VOLUME 102 PART B NUMBER 2 MARCH 1955

The Proceedings

OF

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871: INCORPORATED BY ROYAL CHARTER 1921

PART B

RADIO AND ELECTRONIC ENGINEERING
(INCLUDING COMMUNICATION ENGINEERING)

SAVOY PLACE · LONDON W.C.2

Price Seven Shillings and Sixpence

The Institution of Electrical Engineers

FOUNDED 1871
INCORPORATED BY ROYAL CHARTER 1921

PATRON: HER MAJESTY THE QUEEN

COUNCIL 1954-55

President
J. ECCLES, C.B.E., B.Sc.

Past-Presidents

SIR JAMES SWINBURNE, BART., F.R.S.
W. H. ECCLES, D.Sc., F.R.S.
THE RT. HON. THE EARL OF MOUNT EDGCUMBE, T.D.
J. M. DONALDSON, M.C.
PROF. E. W. MARCHANT, D.Sc.
P. V. HUNTER, C.B.E.
H. T. YOUNG.
SIR GEORGE LEE, O.B.E., M.C.
SIR ARTHUR P. M. FLEMING, C.B.E., D.Eng., LL.D.
J. R. BEARD, C.B.E., M.Sc.
SIR NOEL ASHBRIDGE, B.Sc.(Eng.).
COLONEL SIR A. STANLEY ANGWIN, K.B.E., D.S.O., M.C., T.D., D.Sc.(Eng.).
SIR HARRY RAILING, D.Eng.
P. DUNSHEATH, C.B.E., M.A., D.Sc.(Eng.).
SIR VINCENT Z. DE FERRANTI, M.C.
T. G. N. HALDANE, M.A.
PROF. E. B. MOULLIN, M.A., Sc.D.
SIR ARCHIBALD J. GILL, B.Sc.(Eng.).
SIR JOHN HACKING.
COLONEL B. H. LEESON, C.B.E., T.D.
H. BISHOP, C.B.E., B.Sc.(Eng.).

Vice-Presidents

T. E. GOLDUP, C.B.E.
S. E. GOODALL, M.Sc.(Eng.).
WILLIS JACKSON, D.Sc., D.Phil., F.R.S.
SIR GEORGE H. NELSON.
SIR W. GORDON RADLEY, C.B.E., Ph.D.(Eng.).

Honorary Treasurer
H. W. GRIMMITT.

Ordinary Members of Council

J. BENNETT.
A. R. COOPER.
A. T. CRAWFORD, B.Sc.
C. DANNATT, O.B.E., D.Sc.
B. DONKIN, B.A.
O. W. HUMPHREYS, B.Sc.
C. R. KING, C.B.E.
H. R. L. LAMONT, Ph.D., M.A., B.Sc.
F. J. LANE, O.B.E., M.Sc.
G. S. C. LUCAS, O.B.E.
G. LYON, M.Sc.(Eng.).
SIR HAMISH D. MACLAREN, K.B.E., C.B., D.F.C., LL.D., B.Sc.
A. H. MUMFORD, O.B.E., B.Sc.(Eng.).
W. F. PARKER.
PROF. M. G. SAY, Ph.D., M.Sc.
R. L. SMITH-ROSE, C.B.E., D.Sc., Ph.D.
G. O. WATSON.
J. WESTCOTT, B.Sc.(Eng.), Ph.D.
E. L. E. WHEATCROFT, M.A.
R. T. B. WYNN, C.B.E., M.A.

Chairmen and Past-Chairmen of Sections

Measurements:
M. WHITEHEAD.
*J. F. COALES, O.B.E., M.A.

Radio:
C. W. OATLEY, M.A., M.Sc.
*J. A. SMALE, C.B.E., A.F.C., B.Sc.

Supply:
J. D. PEATTIE, B.Sc.
*L. G. BRAZIER, Ph.D., B.Sc.

Utilization:
J. I. BERNARD, B.Sc.Tech.
*B. L. METCALF, B.Sc.(Eng.).

Chairmen and Past-Chairmen of Local Centres

East Midland Centre:
J. H. MITCHELL, B.Sc., Ph.D.
*C. D. WILKINSON.

North Midland Centre:
W. A. CROCKER.
*G. CATON.

North-Western Centre:
PROF. E. BRADSHAW, M.B.E., M.Sc.Tech., Ph.D.
*H. WEST.

Scottish Centre:
J. S. HASTIE, B.Sc.(Eng.).
*C. H. A. COLLYNS.

Mersey and North Wales Centre:
P. R. DUNN, B.Sc.
*T. COATES, M.Eng.

North-Eastern Centre:
G. W. B. MITCHELL, B.A.
*H. ESTHER, B.Eng.

Northern Ireland Centre:
MAJOR P. L. BARKER, B.Sc.
*J. R. W. MURLAND, B.Sc.(Eng.).

South Midland Centre:
A. R. BLANDFORD.
*H. J. GIBSON, B.Sc.

Southern Centre:
E. A. LOGAN, M.Sc.
*COMDR.(L) C. V. ROBINSON, R.N., O.B.E.

Western Centre:
A. N. IRENS.
*J. VAUGHAN HARRIES.

* Past-Chairman.

RADIO SECTION COMMITTEE 1954-55

Chairman
C. W. OATLEY, M.A., M.Sc.

Vice-Chairmen
H. STANESBY.
R. C. G. WILLIAMS, Ph.D., B.Sc.(Eng.).

Past-Chairmen
J. A. SMALE, C.B.E., A.F.C., B.Sc.
E. C. S. MEGAW, M.B.E., D.Sc.

Ordinary Members of Committee

PROF. H. E. M. BARLOW, Ph.D., B.Sc.(Eng.).
N. R. BLIGH, B.Sc.(Eng.).
G. G. MACFARLANE, Dr.Ing., B.Sc.
B. N. MACLARTY, O.B.E.
J. H. H. MERRIMAN, M.Sc.
BRIG. E. J. H. MOPPETT.
H. PAGE, M.Sc.
A. R. A. RENDALL, Ph.D., B.Sc.
N. C. ROBERTSON, C.M.G., M.B.E.
L. RUSHFORTH, M.B.E., B.Sc.
J. A. SAXTON, D.Sc., Ph.D.
A. M. THORNTON, B.Sc.

And

The President (*ex officio*).
The Chairman of the Papers Committee.
A. H. MUMFORD, O.B.E., B.Sc.(Eng.) (representing the Council).
G. E. MIDDLETON, M.A. (representing the Cambridge Radio Group).
J. MOIR (representing the South Midland Radio Group).
D. H. THOMAS, M.Sc.Tech., B.Sc.(Eng.) (representing the North-Eastern Radio and Measurements Group).

The following nominees of Government Departments:
Admiralty: CAPTAIN(L) G. C. F. WHITAKER, R.N.
Air Ministry: AIR COMMODORE G. H. RANDLE, B.A.
Department of Scientific and Industrial Research: A. F. WILKINS, O.B.E., M.Sc.Tech.
Ministry of Supply: F. S. BARTON, C.B.E., M.A., B.Sc.
Post Office: CAPTAIN C. F. BOOTH, O.B.E.
War Office: COL. J. H. E. PANTON, O.B.E.

Secretary
W. K. BRASHER, C.B.E., M.A., M.I.E.E.

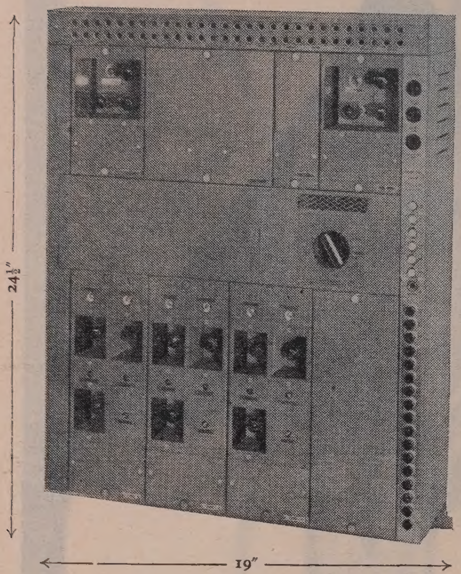
Deputy Secretary
F. JERVIS SMITH, M.I.E.E.

Assistant Secretary
F. C. HARRIS.

Editor-in-Chief
G. E. WILLIAMS, B.Sc.(Eng.), M.I.E.E.

SKILLMAN

3 CHANNEL OPEN WIRE SYSTEM



TEST JACKS

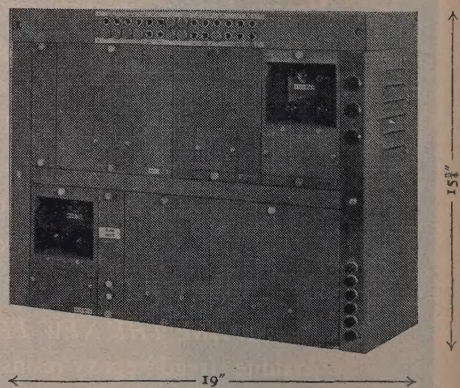
LINE EQUIPMENT

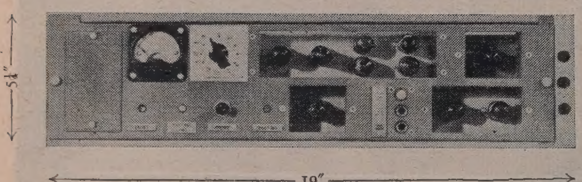
CHANNEL FILTERS

CHANNEL UNITS

SPACE FOR A 4TH CHANNEL

INTERMEDIATE REPEATER





ELECTRONIC AUTOMATIC GAIN
REGULATION FOR COMPENSATING
LINE ATTENUATION VARIATIONS

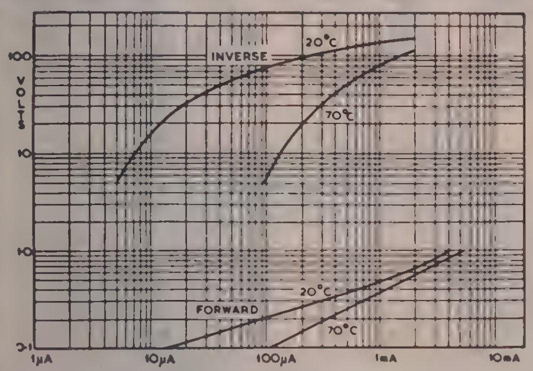
THIS IS NO NEW SYSTEM

ORDERS ALREADY RECEIVED FOR

OVER £250,000

GERMANIUM RECTIFIERS

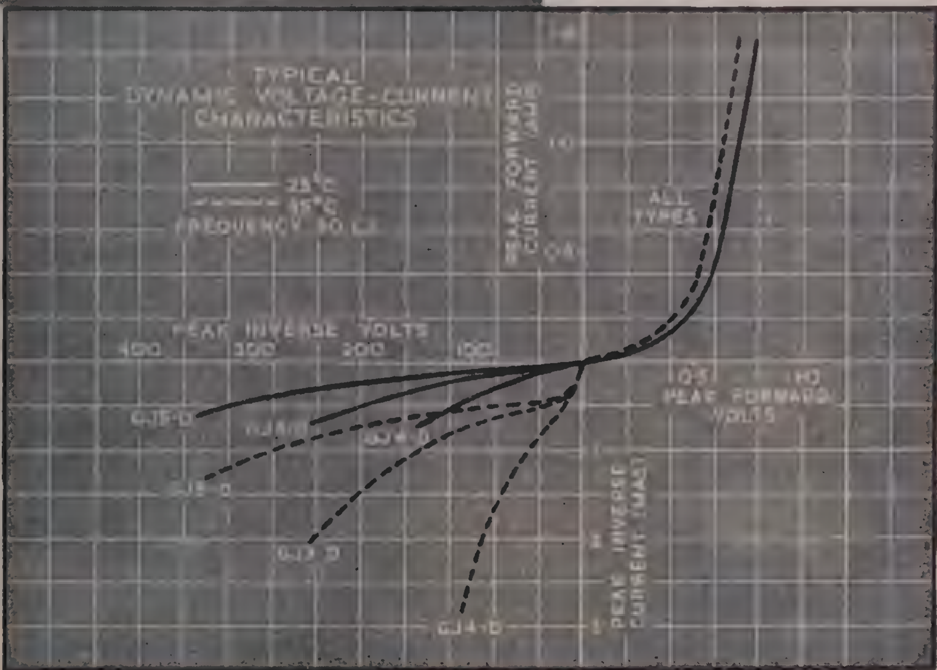
Compact • Reliable • Inexpensive • Efficient



TYPICAL CHARACTERISTIC CURVES

POINT CONTACT RECTIFIERS

A wide range is available, covering applications for television and VHF radio receivers, electronic computors measuring instruments, etc.



JUNCTION RECTIFIERS

These new rectifiers combine very low forward resistance with high back resistance, to give almost 100% rectification efficiency. Their small size makes them eminently suitable for telecommunication circuits, blocking relays, instrumentation applications, and miniature power packs.

BRITISH THOMSON-HOUSTON

THE BRITISH THOMSON-HOUSTON COMPANY LIMITED · RUGBY · ENGLAND

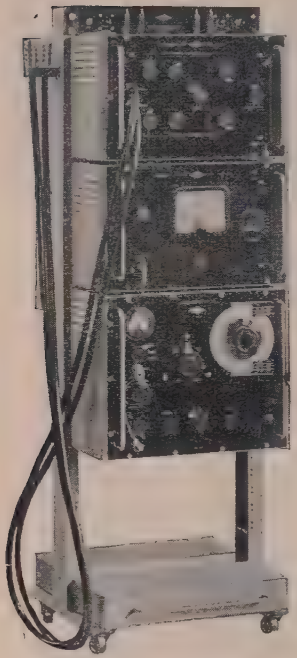
Member of the AEI group of companies

A4850

PHASE MEASURING EQUIPMENT

TYPE RX103

AIRMEC



Frequency Range:	50kc/s–M20c/s.
Minimum Input Level:	60db below 1 volt.
Maximum Input Level:	27db below 1 volt.
Maximum Level Difference:	39db between the two points at which measurements are to be made.
Accuracy:	Attenuation Measurement ±0·5db. Phase Measurement ±3°

THIS equipment has been developed and manufactured to a General Post Office Research Branch design. It was primarily intended for the measurement of the loop phase-shift and gain of feedback repeaters over the frequency range 50 kc/s–20 Mc/s, but it is equally suitable for the measurement of these quantities in amplifiers, filters, equalizers and other four terminal networks.

Full details of this or any other Airmec Instrument will be forwarded gladly upon request

AIRMEC LIMITED

HIGH WYCOMBE BUCKINGHAMSHIRE ENGLAND

Tel.: High Wycombe 2060. *Cables:* Airmec, High Wycombe

Permanent Magnets
for the
Radio & T.V. Industry

The very latest techniques in permanent magnet manufacture can effect real economies in your production and at the same time contribute further to efficient operation of your products. Consult our Technical Staff at an early stage of design to be certain that you are taking advantage of recent progress.

Send for literature:
Pub. PM108—Loud Speaker Magnets.
Pub. PM112—Permanent Magnets for Television Receivers.

Made by the makers of 'Eclipse' Permanent Magnet Chucks

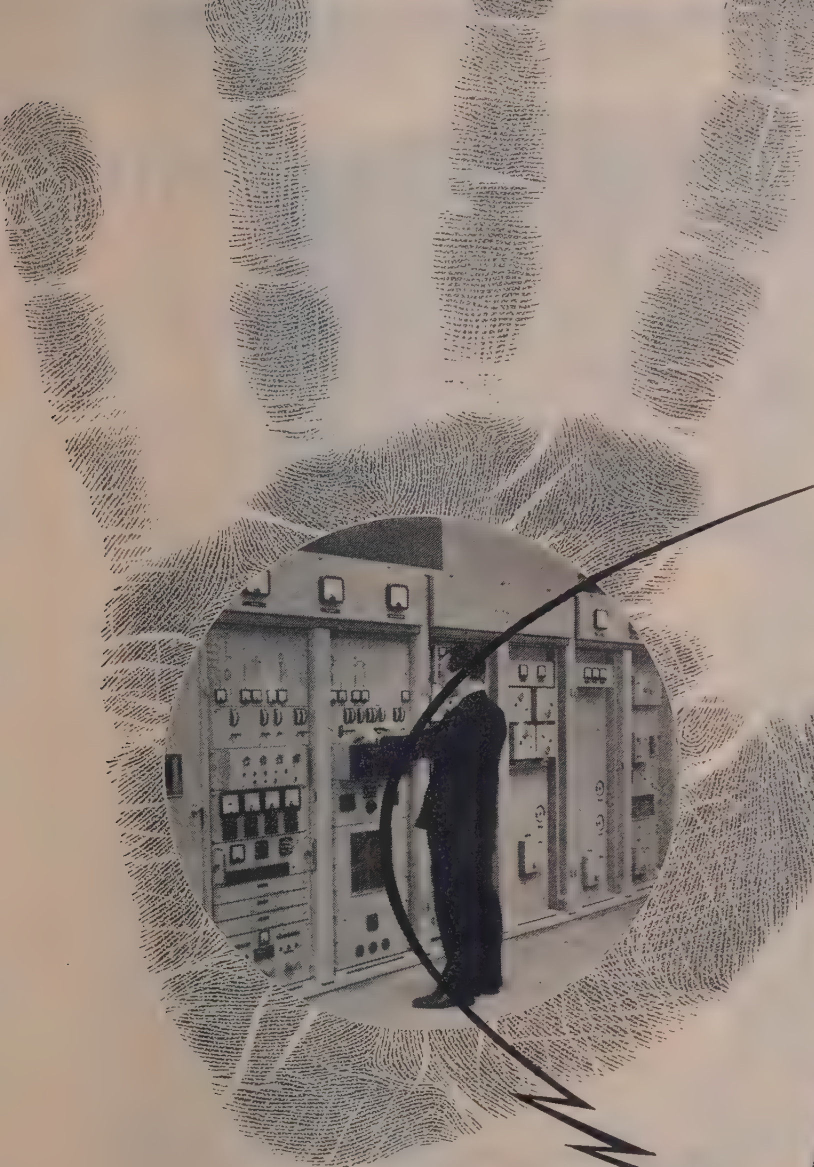
Eclipse PERMANENT MAGNETS

JAMES NEILL & CO. (SHEFFIELD) LTD., SHEFFIELD, 11 ENGLAND

MS

Life line of communication . . .

World wide radio-communication began with Marconi's Transatlantic messages in 1901. Since then Marconi research and development have been behind every major advance in technique. Marconi equipment today, operating at all frequencies, covers a very wide field of both long and short range radio/telegraph and radio/telephone requirements. Marconi VHF multi-channel equipment can provide for as many as 48 telephone channels and is largely superseding land line or cable routes on grounds of efficiency, economy, ease of installation and maintenance.



COMPLETE RADIO/TELEPHONE

AND RADIO/TELEGRAPH

SYSTEMS AND EQUIPMENT

MARCONI

COMPLETE COMMUNICATION SYSTEMS

Surveyed, Planned, Installed, Maintained

MARCONI'S WIRELESS TELEGRAPH COMPANY LIMITED · CHELMSFORD · ESSEX

LG I

Marconi Surveying Service

LEFT. Balloon operations on the Ipoh-Telok Anson route in Malaya.

RIGHT. The mast is up and the motor generator is running during the survey of the Nigerian multi-channel system.

BELOW. The V.H.F. mobile survey team erect their mast.

Before planning any communication system, and particularly a microwave or V.H.F. multichannel system, a survey of the propagation conditions over the proposed path or area is essential. Similar, but less exhaustive surveys, are also necessary before planning V.H.F. mobile systems. Such surveys are undertaken by Marconi's, one of the very few radio manufacturers who do so. The teams engaged in the work may be called upon to operate in desert, swamp and jungle, over which line and cable routes would be impractical, on windswept moorlands or in densely populated city and surburban areas. Surveys are being, or have already been carried out all over the world, including: Uganda, Kenya, Tanganyika, Nigeria, Gold Coast, Tangier, Azores Norway, Turkey, Greece, Malaya, Ceylon, West Indies, Sweden, and also, of course, in Britain.

Over 80 countries now have Marconi-equipped telegraph and communications services. Many of these are still giving trouble-free service after more than twenty years in operation.

Lifeline of communication

MARCONI

COMPLETE COMMUNICATION SYSTEMS

Surveyed, planned, installed, maintained

MARCONI'S WIRELESS TELEGRAPH CO., LTD., CHELMSFORD, ESSEX

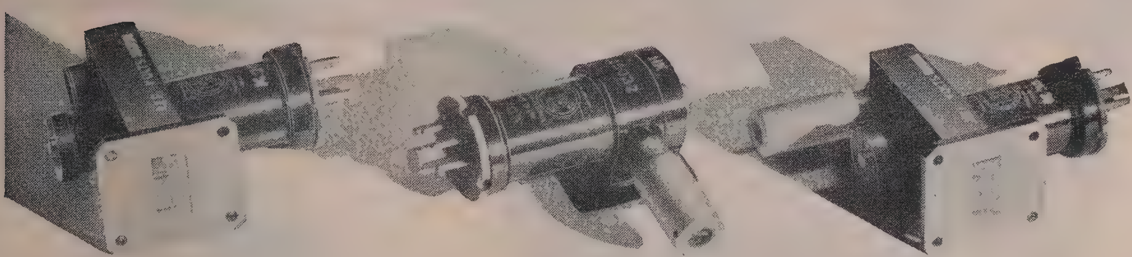
LC 10


ENTRUSTED TO
The Government of the Republic of Haiti has entrusted to the General Electric Company a very considerable extension of their national telephone system. New exchanges and extensions to existing exchanges, local distribution networks, open-wire trunk routes with carrier working, and VHF radio links, all appear in a plan which will give the country a comprehensive, reliable and profitable telephone system.
The plan represents another major step in the steady progress of the country, and fittingly comes in a year that completes a centur
and a half of national independence.
THE GENERAL ELECTRIC CO. LTD. OF ENGLAN
TELEPHONE WORKS, COVENTRY, ENGLAND
Everything for telecommunications by OPEN WIRE LINE, CABLE AND RADIO, Single or Multi-circuit or T.V. link, Short, medium or long haul.
JEREMIE
DEPT. DU SUD
TIBURON.
LES CAYES
Very considerable extension and rehabilitati

KLYSTRONS

by ENGLISH ELECTRIC VALVE COMPANY LIMITED

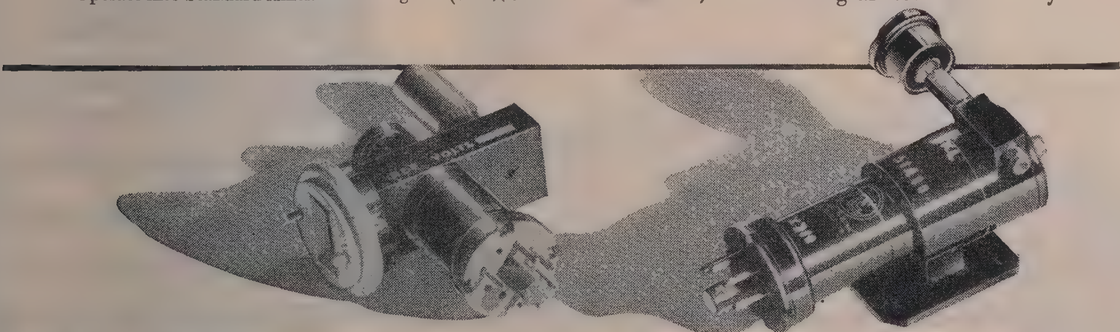
The valves tabulated below are examples from our standard range. The frequency coverage can be varied, within certain limits, to suit the requirements of equipment designers. Further particulars are available on request.



Tube Type	C.V. No.	Minimum Mechanical Frequency Range (Mc/s)	Typical Operation R.F. Power Output (mW)	Typical Operation Electronic Tuning Range (Mc/s)	Type of Tuner
K.300†		9320-9500	25·0	30	Micrometer
K.328†		9550-9700	25·0	30	Micrometer
K.302*	2164	9320-9500	25·0	30	Micrometer
K.305*	2263	9250-9500	25·0	30	External Pin
K.312*	2273	9430-9650	25·0	30	Micrometer
K.313*		9645-9775	25·0	30	External Pin
K.335*	2343	9555-9685	25·0	30	Micrometer
K.308*	2282	8800-8900	30·0	30	Micrometer
K.315*		9105-9205	30·0	30	Micrometer
K.317*		8200-8300	30·0	30	Micrometer
K.311*		8500-9500	40·0	25	Shaft
K.324*	2304	9000-10000	40·0	25	Shaft

† Operate into Standard British Waveguide (1·0″ × 0·5″ Inside Dimensions).
* Operate into Standard American Waveguide (0·9″ × 0·4″ Inside Dimensions).

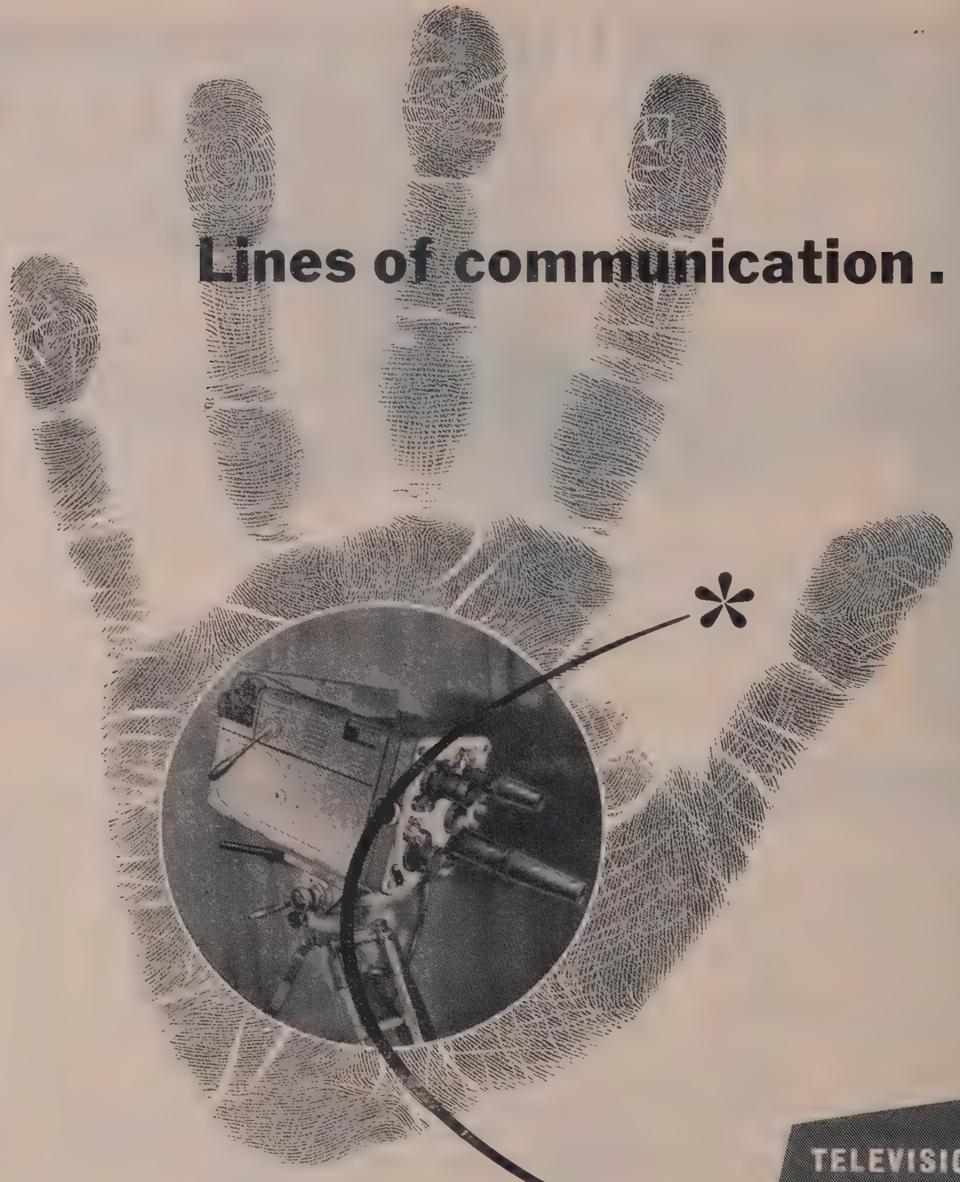
All valves are supplied with an integral resonant cavity.



ENGLISH ELECTRIC VALVE CO. LTD.

Waterhouse Lane, Chelmsford
Telephone: Chelmsford 3491

Lines of communication . . .



In June 1920 the first advertised sound radio programme was broadcast from the Marconi transmitter at Chelmsford. To-day Marconi high or medium power transmitters and high power aerials are installed in every one of the B.B.C's television transmitter stations and Marconi television systems are being supplied to countries in North and South America, Europe and Asia.

TELEVISION CAMERAS
STUDIO EQUIPMENT
RADIO LINKS AND TRANSMITTERS
AERIAL SYSTEMS
COLOUR TELEVISION
INDUSTRIAL TELEVISION

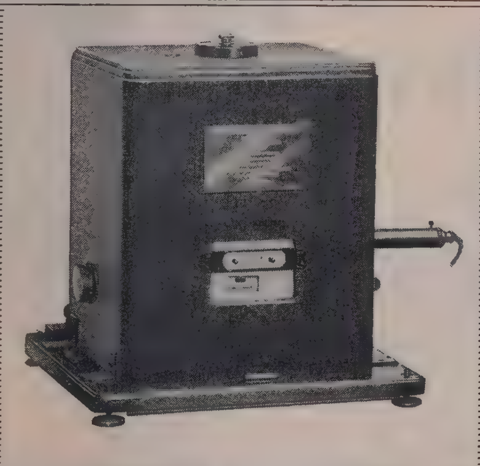
MARCONI

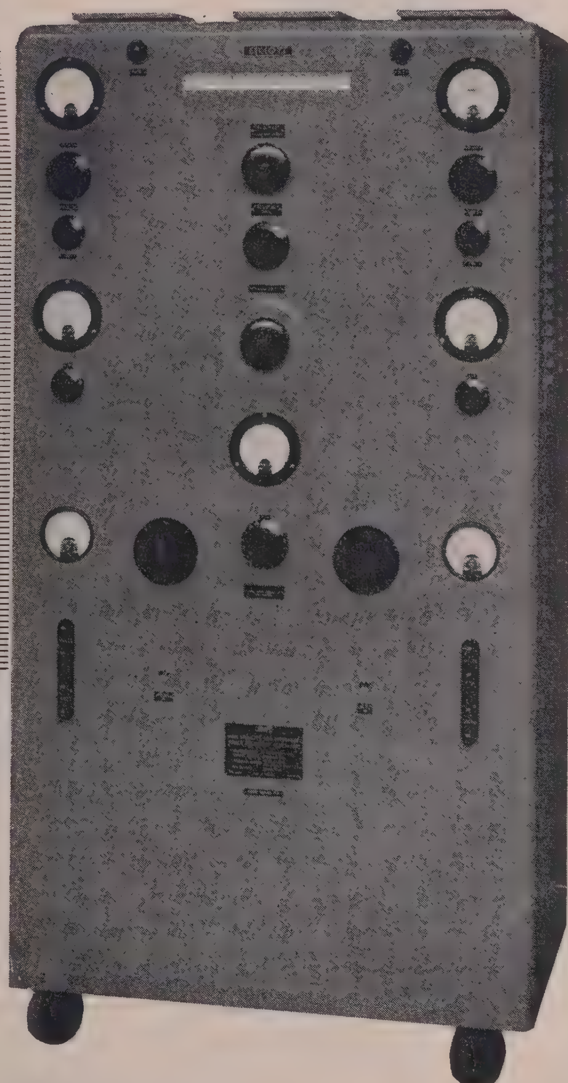
Complete Television and Sound Broadcasting Systems

MARCONI'S WIRELESS TELEGRAPH COMPANY LIMITED · CHELMSFORD · ESSEX

LG 3

ELLIOTT SOLVE THE PROBLEM OF PRECISE MEASUREMENT IN LABORATORIES & TEST STATIONS





ELLIOTT A.C./D.C. COMPARATOR

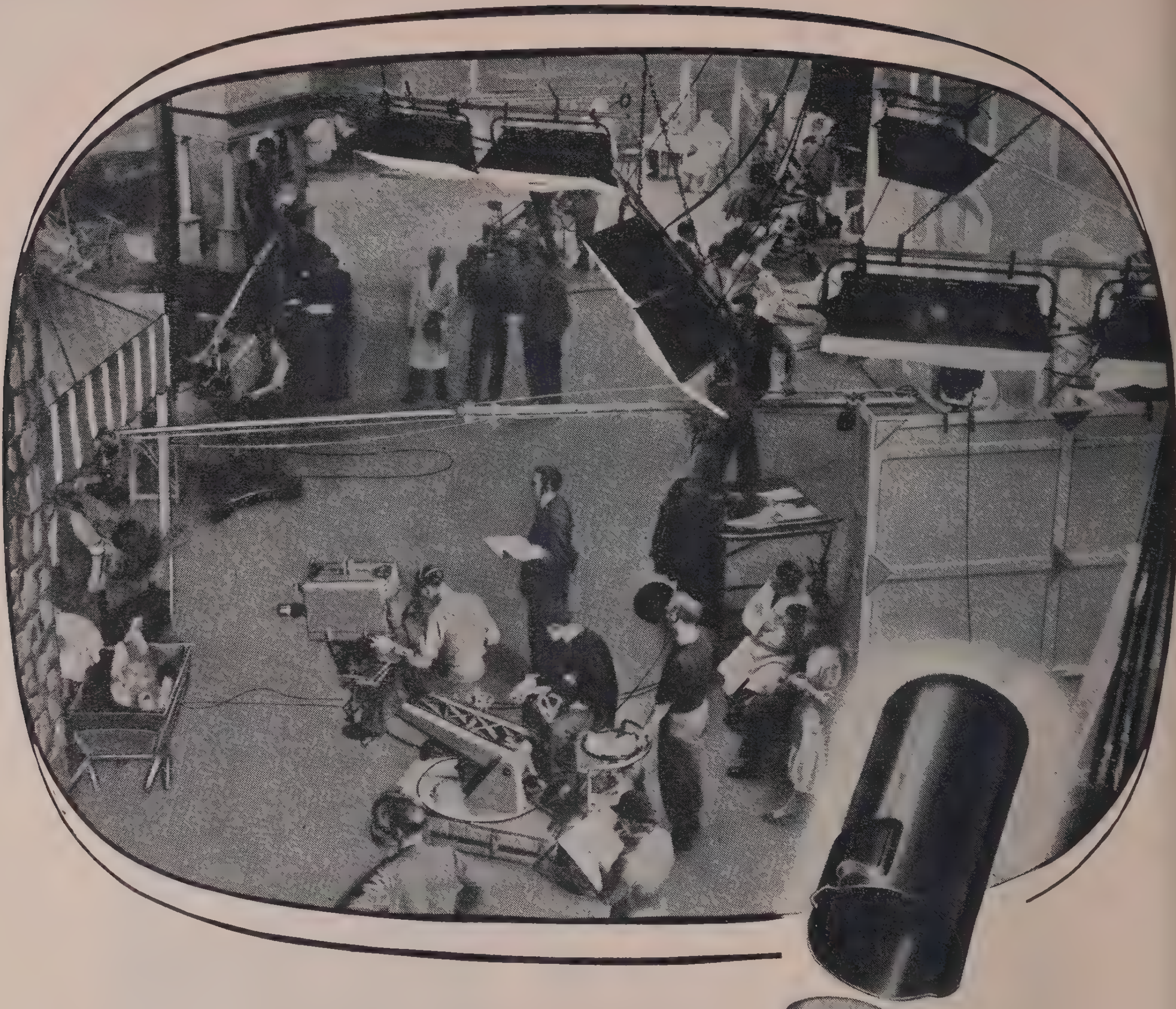
The most accurate commercial means yet devised for the precise measurement of volts, amps and watts at power frequencies. Due to its inherent stability and unique method of self-standardisation this comparator has achieved world acceptance as a means of testing precision instruments.

ELLIOTT HIGH STABILITY POWER GENERATOR

Wherever accurate measurement is required, this generator provides means for testing on A.C. with an accuracy comparable with that obtainable on D.C. It enables the high accuracy of the ELLIOTT A.C./D.C. Comparator to be used to full advantage for the calibration of standard A.C. Ammeters, Voltmeters and Wattmeters. For precise measurements involving the use of electrical standards at frequencies up to 2,500 c.p.s. this generator provides by far the most economical source of supply.

Send for full details of Elliott Instrumentation for the Standards Room.

ELLIOTT BROTHERS (LONDON) LTD., CENTURY WORKS, LONDON, S.E.13 • TELEPHONE: TIDEWAY 3232



Why Ediswan Clix P.T.F.E. Valveholders are widely used in B.B.C. Television equipment

Large quantities of Ediswan Clix P.T.F.E. Valveholders are used in B.B.C. Television equipment.
Only the combination of the finest insulation—P.T.F.E., the most efficient contact material—Berylium copper—and Ediswan Clix design and manufacture can match the requirements of efficiency and reliability in this and all other stringent valveholder applications.
Ediswan Clix P.T.F.E. Valveholders are fully type approved for Services Grade 1, Class 1 conditions. Full details of these valveholders and other components in the Ediswan range are given in catalogue CR. 1681. Manufacturers and Development Groups may have a copy on request.

EDISWAN

CLIX

RADIO, TELEVISION & ELECTRONIC COMPONENTS

THE EDISON SWAN ELECTRIC COMPANY LIMITED, *Member of the A.E.I. Group of Companies*
155 Charing Cross Road, London, W.C.2 and Branches. *Telephone: Gerrard* 8660. *Telegrams: Ediswan Westcent, London*
CR3 Radio Components Sales Office: 21, Bruton Street, London, W.1. *Telephone: Mayfair* 5543

THE PROCEEDINGS OF
THE INSTITUTION OF ELECTRICAL ENGINEERS

EDITED UNDER THE SUPERINTENDENCE OF W. K. BRASHER, C.B.E., M.A., M.I.E.E., SECRETARY

VOL. 102. PART B. No. 2. MARCH 1955

621.315.2.029.4(261)

Paper No. 1741
Sept. 1954

A TRANSATLANTIC TELEPHONE CABLE

By MERVIN J. KELLY, Ph.D., D.Eng., D.Sc., LL.D., Fellow, American I.E.E., SIR GORDON RADLEY, C.B.E., Ph.D.(Eng.), Vice-President, G. W. GILMAN, M.S., Member, American I.E.E., and R. J. HALSEY, B.Sc.(Eng.), Member.

(*The paper was presented jointly to* THE INSTITUTION *and* THE AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS. *It was first received by* THE INSTITUTION *28th June, and in revised form* 22nd July, 1954. *It was published by* THE INSTITUTION *in September*, 1954, *and was read before* THE INSTITUTION 4*th November*, 1954, *the* WESTERN CENTRE 17*th January*, *the* NORTH-WESTERN CENTRE 1*st February*, *the* CAMBRIDGE RADIO GROUP 8*th February*, *and the* NORTH STAFFORDSHIRE SUB-CENTRE 14*th February*, 1955.

SUMMARY

Plans have been announced for the laying of the first transatlantic telephone cable system to link the United Kingdom, Canada and the United States, with a target date for completion in 1956. The paper traces the history of communication across the North Atlantic and discusses the inadequacy of radio circuits to satisfy the requirements of traffic growth. A general description is then given of the cable system, which will provide 36 telephone circuits across the Atlantic and 60 between Newfoundland and Novia Scotia. The project has been made possible only by the development of submerged repeaters containing long-life valves and other components, and these are described, as is the American experimental work which led to the development of a flexible repeater housing which could be laid at depths of 2 miles without risk of damage to itself or to the cable. Fifty-two of these repeaters will be used in each of two one-way cables nearly 2 000 nautical miles long, connecting Newfoundland with Scotland. Newfoundland will be connected to Nova Scotia by a single cable containing 16 both-way repeaters of British design evolved from those which have been extensively used for schemes in home waters.

(1) INTRODUCTION

In December, 1953, plans were announced by the American Telephone and Telegraph Company, the British Post Office and the Canadian Overseas Telecommunication Corporation for the laying of the first transatlantic telephone cable system. The system will provide capacity for 36 high-grade telephone circuits between the United Kingdom, Canada and the United States, and plans contemplate commercial operation at the end of 1956.

In scope the project far exceeds anything that has been attempted hitherto; the two cables on the main crossing will each be nearly 2 000 nautical miles* (n.m.) long and each will contain 52 submerged repeaters, laid at depths of up to 2 000 fathoms, whereas at the time of writing the longest submarine telephone cable is less than 200 n.m. long and no cable contains more than four repeaters. The American Telephone and Telegraph Company and the British Post Office have joint engineering responsibility for the project, which is the outcome of studies and experiments started many years ago. It will represent a major step forward in the evolution of the transatlantic telephone service which began in 1927 with the establishment of one long-wave radio circuit between London and New York.

* One telegraph nautical mile or naut = 6 087 ft = $1\cdot153$ statute miles.

Dr. Kelly and Mr. Gilman are with the Bell Telephone Laboratories.
Sir Gordon Radley and Mr. Halsey are with the British Post Office.

(2) THE NORTH ATLANTIC ROUTE

(2.1) The Early Telegraph Cables

Historically, the North Atlantic route has been the proving ground of most major advances in long-distance oversea communication. The first attempt to establish telegraphic communication between Europe and North America was made by cable in 1857–58, but the cable failed in October, 1858, after a few weeks of service. In 1865 the route once more became the scene of activity, but it was not until the following year that the hazardous operation of laying a cable between Valencia, Ireland, and Trinity Bay, Newfoundland, was accomplished successfully and the North Atlantic route opened to electrical communication.

The successful completion of the first transatlantic telegraph cable was due to the foresight, courage and determination of Cyrus W. Field, an American, coupled with an outstanding British technical accomplishment of its day. The communication capacity of the cable was, however, very limited. Dr. Buckley noted in the 1942 Kelvin Lecture[1] that, even with the improvements in terminal apparatus and operating techniques introduced by Sir William Thomson (later Lord Kelvin) and his associates, the first transatlantic cable operated at a speed of only 3 words/min. This required a bandwidth of perhaps $1\cdot5$ c/s.

To meet the increasing demand for rapid communication across the North Atlantic which arose after the first cable went into operation, more cables were laid, and there are now some 20 in operation on this route.[2] The best available scientific talent—particularly in the United Kingdom—was devoted to increasing the efficiency of the early cables, and late nineteenth century research on submarine telegraph cables and their terminal apparatus may well be one of the first examples of the large-scale effectiveness of applied research. It was not until 1924, however, that the deep-sea cable art took a new turn with the application of Heaviside's earlier suggestion for increasing cable efficiency by means of continuous magnetic loading, implemented through the important discovery of Permalloy.[3] Telegraph cables continuously loaded with Permalloy were laid in the

Atlantic and in the Pacific; the fastest of these cables had about four times the communication capacity of a non-loaded cable of the same size and length and provided an effective bandwidth of about 100 c/s. By comparison the new telephone cable will transmit a bandwidth of 144 000 c/s.

(2.2) Radio Circuits

In December, 1901, Marconi successfully received at St. John's, Newfoundland, radio signals transmitted from Poldhu in Cornwall, and radio entered the field as a competitor to cable on the North Atlantic route. A commercial radiotelegraph service began in 1908.

The next important event in transatlantic communication was the establishment of a commercial radiotelephone service between London and New York in January, 1927, when the long-wave radio transmitting stations at Rugby and Rocky Point were brought into operation. The circuit used the single-sideband suppressed-carrier technique which has since found universal application for both radio and cable telephone circuits. Another important event was the opening of the short-wave (high-frequency) radio spectrum to long-distance communication. Short waves were introduced on the transatlantic route in 1928, and there are, at the time of writing, 14 short-wave radiotelephone circuits between the United Kingdom and the United States and two between the United Kingdom and Canada. Both events were to have an indirect but important effect upon the subsequent development of submarine cables. The availability in the early 1930's of comparatively wide frequency bands for short-wave radiocommunication, the economic depression of that period and other factors combined to defer whatever plans there may have been to lay additional telegraph cables across the North Atlantic. On the other hand, recognition of the technical limitations of radio for transatlantic telephony soon led to studies of the feasibility of a North Atlantic submarine telephone cable.

The advantages of short-wave operation have been

(*a*) Comparatively low-cost installations.
(*b*) Flexibility, i.e. ease of establishing, taking down and re-routing circuits.
(*c*) Direct access from country to country.

With respect to the last point, however, it is of interest to note that arrangements have come into effect to route certain transatlantic traffic through radio-relay stations favourably located to avoid zones where ionospheric conditions are most frequently disturbed. On the other hand, the following disadvantages of short waves are familiar:

(*a*) Although, when ionospheric conditions are stable, short-wave operation furnishes satisfactory transmission, no way has been found to provide day-to-day continuity and reliability comparable to that of good wire lines.
(*b*) The deficiencies of short-wave operation are particularly marked on the North Atlantic route, which is exposed to severe ionospheric disturbances.*
(*c*) The short-wave radio spectrum has with time become overpopulated. Interference between services is an increasingly serious problem and, even when full allowance is made for the development of multi-channel working on a single transmitter, the number of transatlantic circuits cannot be increased very greatly.

The original long-wave radiotelephone circuit, which is still in use, renders an important service in providing limited communication facilities when the short-wave circuits are not serviceable. However, the entire long-wave spectrum has a frequency bandwidth of less than 100 kc/s and, since this must be shared by the world's communication services (telephone and telegraph) that need long waves, long-wave radio cannot be regarded as more than an adjunct for transatlantic telephony.

* The past 25 years have been characterized by three periods of very severe disturbance. The first period covered about three years around 1930. The second, more moderate in effect and of shorter duration, centred on 1944, and the last, which is particularly serious, started in 1950 and is still in progress.

(3) EARLY AMERICAN WORK

In 1919 a small American group was organized under Dr. Buckley to study how to increase the transmission capacity of long deep-water submarine cables; this was in accordance with the emphasis in all American submarine cable work on the problems of deep-water cable technology.

The first work of this group was with short lengths of telephone cables at medium depths. In 1921 three iron-wire (Krarup)-loaded cables were laid between Havana and Key West, a distance of just over 100 n.m. These were the first cables made with an outer copper return conductor giving a coaxial structure. In deep-sea telegraph-cable practice the custom had been to depend upon the sea itself and the armour wires of the cables as a return medium, but this is not satisfactory at telephone frequencies. Each of the three Key West–Havana cables provided sufficient frequency bandwidth to accommodate one telephone and three carrier telegraph circuits in addition to a d.c. telegraph circuit. Two years later two cables were put into operation over the 23 n.m. route between the California shore and Catalina Island, and these were ultimately operated at frequencies up to 35 kc/s. The Key West–Havana and Catalina cables were operated as both-way cables, that is each cable was used for both directions of transmission simultaneously. Separation of the two directions of transmission was by means of balancing circuits at the terminals.

There soon followed the developments in long deep-water high-speed submarine telegraph cables mentioned earlier.

By the late 1920's developments in submarine-cable technology and in magnetic and insulating materials reached a point where a single-circuit telephone cable across the North Atlantic between Ireland and Newfoundland (a distance of 1 800 n.m.) seemed feasible. A specific proposal to this effect was formulated by the Bell Telephone System in 1928 and discussed with the British Post Office. The proposal called for laying a single continuously-loaded non-repeatered coaxial cable, operating at a top frequency of about 3 800 c/s and providing a single telephone circuit. The central conductor was to have been loaded with wrappings of tape made of a newly-developed magnetic material, Perminvar. The insulation of the cable was to have been Paragutta, also a new development.

As a check on the estimates of the performance of this design, and to determine whether it could be successfully laid and recovered, a 20-mile section was manufactured in Germany and laid in deep water in the Bay of Biscay in 1930. This was subsequently recovered and relaid in shallow water. A length of non-loaded cable with Paragutta insulation was laid the same year between Key West and Havana. The projected cable was never laid across the Atlantic, principally because of economic conditions. However, studies and experiments in the laboratory and at sea, looking towards a new and better cable, continued during the depression following 1930.

During the laying and recovering operations in the Bay of Biscay in 1930, "kinking" of the cable and buckling of the conductors occurred on several occasions. This extreme local deformation was caused by unwrapping of the armour wires at points along the cable; it happens in the following way. As the cable passes aft of the ship, variations in tension cause twisting and untwisting of the cable structure; this twisting is harmless if it is distributed along the length of the cable but buckling can occur if it is localized for any reason. Stopping and starting of the cable ship or the existence of any large irregularity along the cable itself are potential causes of damage; for example, severe local distortion was produced artificially during the Bay of Biscay tests when a large board, nicknamed the "barn door," was lashed to the cable.

The structure of the proposed coaxial transatlantic telephone

able was inherently much more complex than that of the conventional non-loaded deep-water telegraph cable. The cable was therefore more subject to damage while being handled at sea and, because of the serious consequences of such damage, great attention was paid to the problem of maintaining the structure of the cable unaltered during laying operations. Laboratory studies led to a design in which the lays of the tape loading, the tapes of the copper inner and outer conductors and of the armour wires were in the same direction. This was a departure from established thinking at the time, but has been carried through in the subsequent cable designs of the Bell Telephone Laboratories.

(4) MULTI-CHANNEL TRANSATLANTIC CABLE PROJECT: BACKGROUND OF AMERICAN EXPERIENCE

By 1928, when the proposal for a transatlantic telephone cable was first formulated, electronic techniques were well established in land-cable practice, although they had not reached the point when serious proposals could be made to lay a transatlantic submarine cable with submerged repeaters lying on the ocean bottom. For this to become feasible, further improvements in technology were required, notably the use of amplifiers stabilized by negative feedback and pentode valves with equipotential and uncombined cathodes. These developments, which were to make long-distance multi-circuit carrier telephony a reality within the next ten years, together with still greater reliability, were needed to make the transatlantic plan practicable. Until these improvements became available, the only kind of repeatered transatlantic system that could be imagined was one in which the repeaters would be mounted on moored platforms at sea or would be in submerged buoys, fed with power from local batteries and visited at intervals for maintenance and repair. Systems of this kind were considered and discarded as impracticable.

By 1932 great advances were being made and the first experiments on long-distance broad-band systems for use on land were well under way. While the solution of many of the submarine-cable problems was not yet clear, the rapidly developing electronic art created the atmosphere of confidence that made it possible to contemplate seriously the design of a long-distance deep-water cable associated with submerged repeaters in which were long-life valves which would be capable of transmitting a number of carrier telephone channels instead of only one channel.

The broad concept which grew up of the transatlantic cable system was the use of two non-loaded coaxial cables, one for each direction of transmission, into which repeaters would be spliced at regular intervals. Detailed discussion of the relative merits of two cables or one is outside the scope of the paper; suffice it to say that an economic field for the use of twin cables (one-way operation in each) may be on long deep-water routes with traffic rapidly growing beyond the capacity of a single-cable system. In this case, however, the choice was influenced by the need to use repeater housings which would cause the minimum irregularity in the cable structure. As shown in Section 5, both-way repeaters for use in a single-cable system require more accommodation in the repeater housing than do repeaters for one-way operation.

The repeaters required for the transatlantic cable system had to be designed to withstand the shocks of laying and recovery and the pressure of water encountered at the greatest depth on the North Atlantic route. The valves used had to operate at sufficiently low cathode power and anode potential to make it feasible to supply power to the repeaters over the cable from the shore ends at a safe working potential, which was eventually established at roughly 2 000 volts d.c. Finally, the repeaters had to have sufficiently long lives to operate with small likelihood of failure over a period of time (generally taken as 20 years) which would be long enough to make the enterprise as a whole technically sound and economically attractive.

(4.1) Repeater Housing

Early experience of the potential hazards of damage to, or modification of, the structure of the cable during handling led to the objective of a flexible housing, structurally as nearly as possible like the cable and capable of being passed over the drum and sheaves of the cable ship and paid out without slowing up or stopping the vessel or suffering damage. The so-called "snake type" housing for the valves and other repeater elements was developed, and will be described in Section 8.2.

(4.2) Long-Life Valves

The valve finally adopted was a conservatively-designed suppressor-grid pentode whose details were finalized in 1941 and which has not since been changed in any essential way. It is a valve with large cathode area, low cathode temperature (630° C), low anode potential (60 volts), large internal spacing between elements and of generally rugged and shockproof construction. Some of these valves have been under continuous test for more than 13 years, and 18 have been in operation for four years in repeaters on the most recent Key West–Havana cables.[4] All indications are favourable for something better than 20-year life expectancy for North Atlantic service.

(4.3) Components

In addition to the valves, about 60 other circuit elements have to be accommodated inside the submerged repeater housing, including resistors, capacitors, inductors, transformers and crystals. Like the valves, these elements must be designed and fabricated so as to avoid all possible risk of failure in service; they must also meet requirements of ruggedness and reasonably small size. Cost, fortunately, is a less important factor than it is in most other applications.

The failure of any one of the other components could be as serious as that of a valve, and there was evolved in the Bell Telephone Laboratories a 5-step philosophy for their selection and manufacture aimed at achieving as near perfection as humanly possible. The steps are as follows:

(*a*) Draw on the background of experience with similar parts in use in the telephone plant.
(*b*) Supplement experience with extraordinary care in selecting material and with special testing of components during production and after completion.
(*c*) Use an open and simple structure to avoid hidden errors.
(*d*) Select and train personnel for perfection of skill and pride in product.
(*e*) Develop the worker's faculty of self-criticism, so that only those parts are submitted which truly represent his best efforts.

In this way there was incorporated into the development of each component all the knowledge and integrity that American telephone practice and experience provided. Wire-wound resistors, for example, have been made for 50 years, and there are something like 50 million of them in service. This type of resistor was therefore chosen for the submerged repeater in preference to the deposited-carbon resistor of more recent origin. Similarly, experience in the control of materials and quality production and experience in the telephone operating field were drawn upon in the selection of designs of Permalloy-core inductors, and paper and mica capacitors. This experience was supplemented wherever possible by accelerated life-tests.

Care in the control of materials used for the repeater elements has been particularly important. A case in point is the selection of paper for the capacitors. Large quantities of this paper are purchased annually for the manufacture of capacitors for use in

the general telephone plant; the paper is supplied in lots which are submitted to accelerated life-tests. The lots which performed best in an interval of six months were made available to the Laboratories, where every roll was re-inspected, and only those showing the highest life were chosen for the capacitors.

(4.4) Cable

The design of cable finally decided upon consists of a solid-dielectric coaxial structure covered by the usual jute and armour for protection and strength. The composite central copper conductor consists of a central solid wire surrounded by a single layer of spiralling and abutting tapes, carefully formed to fit closely about the wire with a minimum of voids. Such a composite structure has better flexibility and protection against breakage than a solid conductor and has better a.c. and d.c. resistance than a stranded-wire conductor of the same outside diameter. The return conductor is also a flexible composite copper structure consisting of a single layer of abutting spiralling tapes, carefully formed into a tubular configuration. It is covered with an overlapping spiralling copper tape for protection against the teredo worm. The various spiralling layers of tape and armour are applied in the same direction of lay, for reasons stated in Section 3.

The original coaxial-cable designs had used Paragutta as the dielectric, but shortly before the Second World War representatives of the Post Office brought polyethylene (Polythene)—a British development—to the attention of the Bell Telephone Laboratories. On test it proved to have characteristics that made it superior to Paragutta; it had a lower permittivity, greater electric strength, was more uniform in its characteristics, was less pervious to sea water and had mechanical properties that made it more desirable. Polyethylene was soon adopted as a feature of the proposed transatlantic cable.

(4.5) Key West–Havana Cables

During the Second World War the American group concerned with the cable was reduced to the barest minimum necessary to keep the development alive, but a machine to simulate cable-laying operations was constructed and used to test ideas concerning repeater housings.

Upon the cessation of hostilities the decision was taken to install on the Key West–Havana route a pair of cables of the design that had been evolved for transatlantic service. A special air-conditioned repeater assembly shop was set up at the Laboratories and a contract was entered into with a manufacturer for experimental cables. Arrangements were also made for modification to the cable ship *Lord Kelvin*. By 1948, work had progressed to a point which enabled 20n.m. of cable and repeaters, together with special built-in testing facilities (accelerometers) for observing and measuring shocks to the cable during handling, to be laid in depths up to 2n.m. in the Bahamas. During this period of growing activity, methods of cable manufacture, including testing and quality control, were worked out in co-operation with the manufacturer.

This work culminated in May, 1950, with the laying of the Key West–Havana cables[4] (115 and 125n.m.). The cable system includes six repeaters, five of which are on the ocean bottom at depths ranging from 20 fathoms to 950 fathoms. The sixth repeater is located in a cable vault on the shore near the Havana telephone office. These repeaters have been in operation, at the time of writing, for four years without failure and without indication of deterioration.

(5) BRITISH SHALLOW-WATER EXPERIENCE

Many submarine telephone cables connect the United Kingdom to the continent of Europe; they are of a variety of types and lie in water which is nowhere deeper than 300 fathoms. The first coaxial cable of modern type, and generally similar to the cable described in Section 4.4, was laid in 1937 and used Paragutta dielectric. All post-war coaxial cables, however, have been constructed with polyethylene.

British development of submerged repeaters for existing cables has recently been described in a series of papers.[5–7] Experimental work by the Post Office did not start until 1938, but in 1943 a coaxial cable between Anglesey and the Isle of Man was equipped with a shallow-water repeater.[8] In 1946 a repeater was inserted into a cable between Lowestoft and Borkum to provide direct telephone communication to the British Armed Forces in Germany. These were the first uses of submerged repeaters for telephone traffic anywhere in the world. In these two early applications, "go" and "return" speech channels were provided over a single cable and separated on a frequency basis, the repeater being arranged to amplify only the channels transmitted in the high-frequency band.

The Post Office next turned its attention to the development of submerged repeaters suitable for operation in tandem and designed to transmit a supergroup of 60 circuits over existing coaxial cables.[6] At this stage a decision became necessary as to whether British development should continue to be in accordance with a pattern of transmitting "go" and "return" channels over a single cable, which was the established practice with unrepeatered systems, or whether separate cables should be used for the two directions. The use of separate cables for the two directions of transmission meant greater simplicity of equipment, but Post Office experience of the inherent difficulties of submarine-cable maintenance in the North Sea had led to a firm preference for making every cable self-contained in respect of circuits where the traffic did not necessitate more than two or three cables on a route.

It was therefore decided that the repeaters should be designed for two-way transmission. The primary supergroup of 60 speech channels is, by international agreement, assembled in the frequency band 312–552 kc/s, and the system was arranged to transmit this frequency band in one direction and the frequency band 24–264 kc/s in the other. High- and low-pass filters were necessary to separate "go" and "return" channels at each repeater, but these were arranged to feed into a common amplifier.

The directional filters necessary for two-way transmission waste a part of the frequency spectrum, while the added components somewhat increase the probability of premature failure and require additional accommodation in the housings. The loss of frequency spectrum and slightly increased risk of failure were not considered important in comparison with the operational advantages of single-cable systems, and the additional accommodation required was not a serious problem because experience had shown that the large single-ended housings which were in use at the time could be laid and recovered without difficulty in shallow water.

The system evolved, which may include up to ten intermediate repeaters, appears adequate for all routes between the United Kingdom and Europe. It was first installed in 1950 and about 20 repeaters are in operation at the time of writing. Very similar 36-circuit systems have been in operation for four years between the Netherlands and Denmark.[7]

These earlier projects, while profitable and affording valuable experience, did not lead directly to a design suitable for a transatlantic project or for installation elsewhere at ocean depths. In 1948, however, the Post Office began to study the possibility of using submerged repeaters to increase the traffic capacity of the telegraph cables operated by Cable and Wireless, Ltd., and development of a housing suitable for installation in deep water was commenced in collaboration with a manufacturer. The

design was modified after laying tests in 2 700 fathoms in the Bay of Biscay in 1951, which were carried through under adverse weather conditions and were not entirely successful. This housing (described in Section 9.3) is rigid, like the shallow-water unit, but is double-ended and much smaller; it is in line with the cable, is designed to transmit the full laying tension experienced in deep water and is able to rotate freely during the laying operation. Limited tests were carried out in the Mediterranean Sea at depths of 250 fathoms in the autumn of 1953, when lengths of cable including the new repeaters were repeatedly laid and recovered. Twist, tilt and recording devices included in one repeater housing established that the repeater could be handled and laid without significant bumping. It has yet to be established that the design is suitable for use in very deep water, and in particular that the laying operation does not hazard the cable by the insertion of slack which subsequently tightens into kinks.

The Post Office did not commence serious study of the transatlantic problem until 1950, when it set up a committee to report on the future possibilities of repeatered telephone cables on this route. In 1952 it was decided to engineer a new telephone cable of about 300n.m. between Scotland and Scandinavia as a prototype for a system suitable for water of greater depth than that actually on the route, although this is no more than about 200 fathoms, and the requirements of 36 circuits on a single cable could have been met using the proved shallow-water repeater design. The cable will be laid in the autumn of 1954. An innovation in submerged-repeater circuit design will be the use in each repeater of parallel-connected amplifier units with two separate forward paths and a common feedback path. The valves will be of a long-life type developed in the Post Office Research Laboratories for submerged-repeater applications.

The decision to use deep-water-type repeaters for this system was an important step in building up a background of experience which will be utilized in the Newfoundland–Nova Scotia section of the transatlantic system.

(6) DECISION ON TRANSATLANTIC TELEPHONE CABLE PROJECT

Early in 1952 negotiations were once again opened between the American Telephone and Telegraph Company and the British Post Office. Two of the present authors* reported jointly on the status of submerged-repeater development in the United States and in the United Kingdom after independent studies by engineers of the Company and of the Post Office.

At that time the American development had been completed in its basic features for several years. The valves and other components, initially constructed or selected for reliability in service, had been under life observation since about 1940. Cable and repeater designs for a transatlantic project had been subject to stringent laboratory tests and had finally been given the practical trials, already referred to, on the Key West–Havana route. The authors concerned recommended that these designs should be used for the long link between Newfoundland and Scotland, because proved reliability is an essential requirement in a pioneering and costly venture of the difficulty of a transatlantic telephone cable.

There had not been the opportunity to subject the elements in the British development—particularly those which it was proposed to introduce as parts of a deep-water, as distinct from a shallow-water, system—to comparable tests. Nevertheless, it was apparent that, if means for laying and recovering the large repeater housings in deep water without damage to the cable were later developed, the design would provide great flexibility in repeatered cable systems of the future. The authors therefore recommended that the British design should be used for the link between Newfoundland and Nova Scotia where the water is of moderate depth and offered the opportunity for observation of a design based on proved shallow-water techniques. The design, because of its later initiation, permitted of a lower cost per telephone circuit on this section.

* Kelly and Radley.

The Canadian Overseas Telecommunication Corporation was asked to join in the enterprise, because of its interest in oversea communication and of the desire that the project should also improve communication between the United Kingdom and Canada. Administrative and technical discussions took place on both sides of the Atlantic, and agreements were arrived at and formalized in a contract signed on 27th November, 1953. Ownership of the system will be vested in the American Telephone and Telegraph Company, the British Post Office, the Canadian Overseas Telecommunication Corporation and the Eastern Telephone and Telegraph Company. The latter is a subsidiary of the American Telephone and Telegraph Company and will be concerned with the facilities within Canadian territorial limits.

(7) GENERAL DESCRIPTION OF THE COMPLETE SYSTEM

(7.1) Route

A map of the system is shown in Fig. 1, and the route in Newfoundland is shown in greater detail in Fig. 2. The route between Clarenville in Newfoundland and Oban in Scotland has been chosen so as to be well clear of existing telegraph cables, which lie to the south, and also to avoid known "holes" in the ocean bed. Selection of a landing site in Newfoundland was determined by the following considerations:

(*a*) The cables remain in deep water until they enter the north-western arm of Random Sound where damage is not to be anticipated from grounding icebergs.
(*b*) There is no crossing over other cables.
(*c*) Clarenville provides reasonable amenities for the staff who will have to maintain equipment associated with the cables to Scotland and Nova Scotia.

Both the cables between Clarenville and Oban will be approximately 1 950n.m. long and each will be equipped with 52 repeaters. The greatest depth at which a repeater will lie will be approximately 2 300 fathoms (2·6 statute miles), but most will be at depths between 1 200 and 2 000 fathoms.

Three alternative routes connecting Clarenville with Nova Scotia were considered; none was particularly attractive from an engineering standpoint. They were:

(*a*) To leave Clarenville generally in an easterly direction and to lay the cable in a wide sweep round the Avalon Peninsula and in deep water to the south of the Newfoundland Banks.
(*b*) To leave Clarenville generally in a southerly direction, cross the Avalon Isthmus (about five miles wide) into Placentia Bay, cross the Burin Peninsula (about ten miles wide) into Fortune Bay and thence in water of moderate depth to Sydney Mines.
(*c*) To lay the cable overland to Terrenceville at the head of Fortune Bay, thence picking up the same inshore route as in (*b*) above.

The first of these would have involved a cable 600n.m. long crossing a number of existing telegraph cables; the second, while avoiding the main fishing grounds and other cables, would have necessitated landings at four intermediate points. Therefore, after an on-site investigation, the last alternative was chosen. This section thus falls into two parts: to the east about 62 statute miles overland and to the west about 274n.m. in coastal waters at depths up to about 250 fathoms. The single cable will be equipped with 16 repeaters, two of which will be on land.

(7.2) Circuit Capacity

Fig. 3 shows the location of the various frequency bands used for transmission and the location of various pilot tones. The

Fig. 1.—The route across the Atlantic.

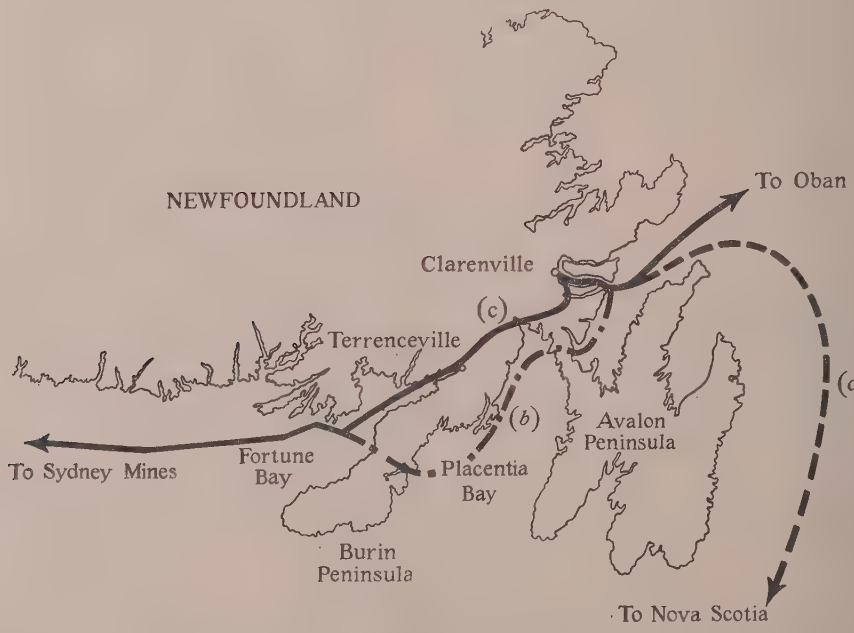


Fig. 2.—The routes in Newfoundland.

cables between Clarenville and Oban will have transmission frequency bands extending from 20 to 164 kc/s and will provide capacity for 36 carrier telephone circuits, with channels at 4-kc/s spacing. The capacity of the cable between Clarenville and Sydney Mines will be somewhat greater, because it is planned to transmit the frequency band 20–260 kc/s from Sydney Mines to Clarenville and the frequency band 312–552 kc/s in the reverse direction. The cable will therefore provide 60 carrier telephone circuits with channels spaced 4 kc/s apart. One of the two 12-channel groups in each direction, not employed for transatlantic service, will be used to provide 12 circuits for "local" telephone services between Newfoundland and the rest of Canada; the other will be held spare for the present.

The junction point between the east and west parts of the system at Clarenville will be the group distribution frame. The transatlantic channels in both parts will appear on this frame as primary groups in the frequency range 60–108 kc/s. From the frame, connection will also be made to the local and spare

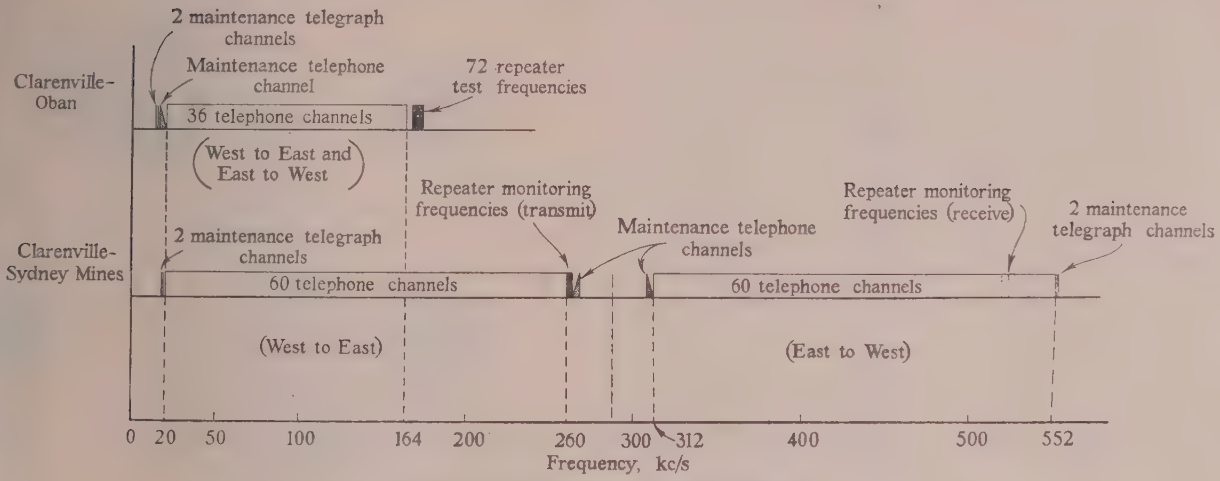


Fig. 3.—Allocation of the frequency bands.

groups. Equipment to bring each of the 12 channels in any one transatlantic group down to audio frequency will be installed, but not normally connected. The availability of such equipment is desirable, in order to enable impairments affecting only one or two channels to be localized to one part of the system. Fig. 4 band below 20 kc/s, as indicated in Fig. 3. Between Clarenville and Sydney Mines the telephone channels will be provided in the "cross-over" frequency band between 260 and 312 kc/s and the telegraph channels below 20 kc/s in one direction and above 552 kc/s in the other direction.

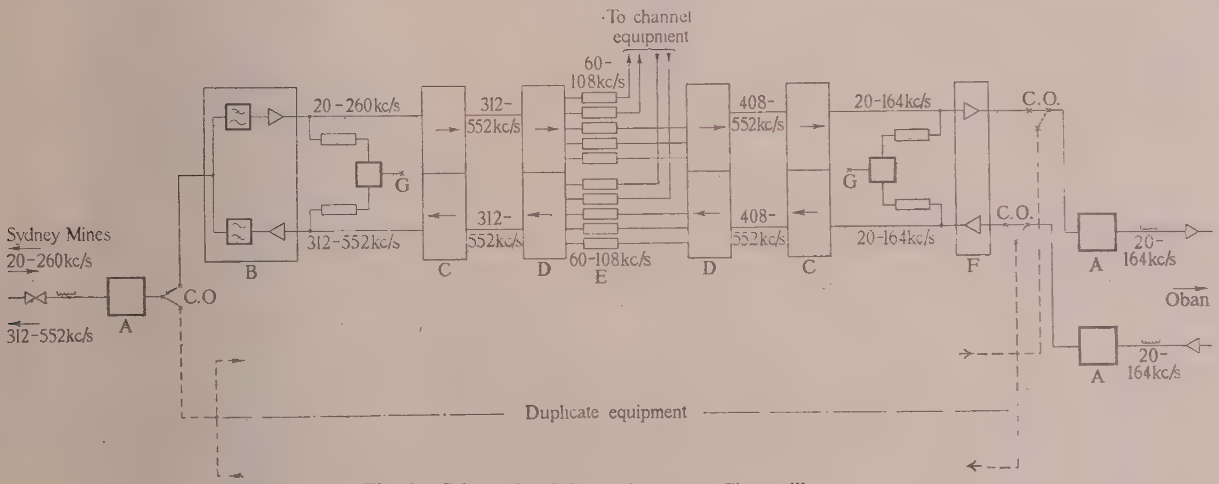


Fig. 4.—Schematic of the equipment at Clarenville.

A. Cable terminating equipment.
B. Terminal filters and amplifiers.
C. Terminal translating equipment.
D. Group translating equipment.
E. Through-group filters.
F. Terminal amplifiers.
G. Maintenance circuits.

C.O. Change-over switch to duplicate equipment.
At Oban and Sydney Mines the schematics are complementary to those of the facing equipments at Clarenville.

is a schematic of the arrangements at Clarenville; the arrangements will be generally similar at Oban and Sydney Mines.

The traffic operating terminals will be in London, New York and Montreal. Twenty-nine telephone circuits will be put into service between London and New York, and six between London and Montreal. Telegraph circuits between London and Montreal will be provided in the upper half of the thirty-sixth telephone circuit.

Telephone and telegraph circuits for maintenance purposes will be provided between Oban and Clarenville in the frequency

(7.3) Overall Performance

Since the transatlantic circuits will serve to connect two extensive telephone networks on opposite sides of the ocean, including connection to the European continent, which will be reached through London, they are being designed to add as little extra loss and other forms of impairment as possible. For example, it has been agreed that the target for the frequency characteristics of the channels between London and New York, or London and Montreal, should be at least as good as that specified by the C.C.I.F. for a 2 500 km international circuit.

Other objectives for the circuits compare favourably with those specified for long-distance circuits wholly on land.* For the transatlantic circuits the "via net loss" will be 0·5 db measured at 800 or 1 000 c/s between London and New York (or Montreal). With a net loss of this order on circuits which will have a one-way transmission time of about 35 millisec, echo control will be necessary, and a study is being made of the best method of applying this. Terminating circuit pads will add 2 db to the loss at New York and 3·5 db at London for calls between the two cities, but when the transatlantic circuits are extended the pads will be switched out of circuit. Thus the normal attenuation on a London–New York transmission will be 6 db (3·5 + 0·5 + 2) or slightly more than the accepted value in regional-centre to regional-centre connections in the United States or Canada. It is an agreed objective to restrict variations from the normal attenuation to a standard deviation of 1·5 db, and this will put a severe requirement on the individual links which may be difficult to meet. In all respects the requirements for the Sydney Mines–Clarenville–Oban cable sections have to be somewhat more stringent than for the overall circuits.

A pilot tone of approximately 84 kc/s† will be injected into each group of 12 channels at a point where the channels appear in the 60–108-kc/s range. Measurements of the received level of the tone will give an indication of the overall attenuation; later the received tone may be used to control automatically the gain of terminal amplifiers on each group. In addition, 92-kc/s pilot tones, similarly injected, will be transmitted over each section to give an indication of the attenuation of the section. These tones will be blocked from passing from one section to another.

Signalling over the cable circuits will be designed to allow co-ordination between the different national long-distance signalling systems at the ends.

(8) CLARENVILLE–OBAN SECTION

As already indicated, two cables will be used between Newfoundland and Scotland, one for each direction of transmission. These and the associated repeaters will be of the type now in operation between Key West and Havana and have been described in some detail in the literature.[4]

(8.1) Cable

The evolution of the cable design has been described in Sections 3 and 4.4. The cable to be used on the transatlantic route will have a somewhat greater thickness of dielectric than that included in the make-up of the Key West–Havana cables and will be 0·620 in in diameter.‡

Manufacture of the cable will demand control of the dimensional tolerances to a greater degree of precision than has been attempted hitherto, because the large number of repeaters in circuit make it of great importance to maintain a fine balance between repeater gain and cable loss. Predetermination of the effects of temperature and pressure on cable attenuation will enable this fine balance to be obtained when the cable is at sea-bottom temperature and pressure. Impedance irregularities must also be avoided, since they affect the gain/frequency characteristics of the repeaters.

As is customary in cable practice, special cable armouring will be used near the shore ends, and the various types of armouring

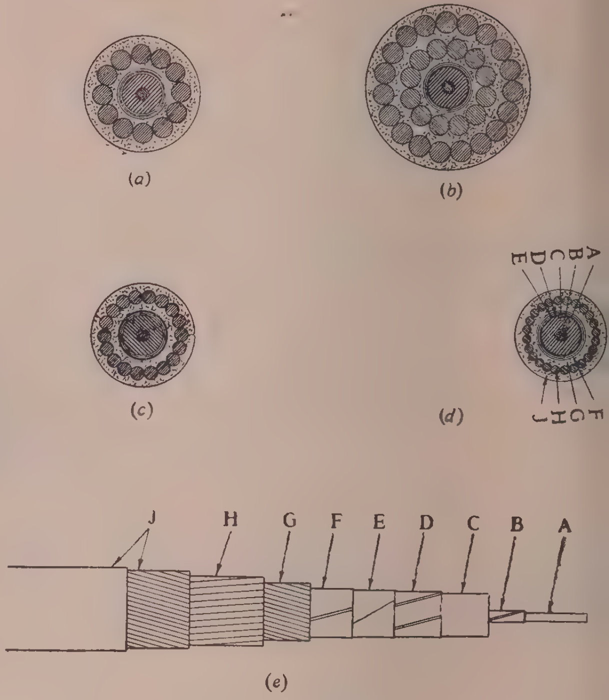


Fig. 5.—Types of cable.

(a) Type A; overall diameter 1·84 in.
(b) Type AA; overall diameter 2·68 in.
(c) Type B; overall diameter 1·40 in.
(d) Type D; overall diameter 1·21 in.
(e) Construction of type-D cable.

A. Centre conductor; 0·1318 in diameter copper.
B. Three 0·0145 in copper surround tapes.
C. Polythene to 0·620 in diameter.
D. Six 0·016 in copper return tapes.
E. 0·003 in overlapped copper teredo tape.
F. Gapped Telconax tape.
G. One serving of cutched jute.
H. Twenty-four 0·086 in diameter high-tensile steel armour wires.
J. Two impregnated-jute servings.

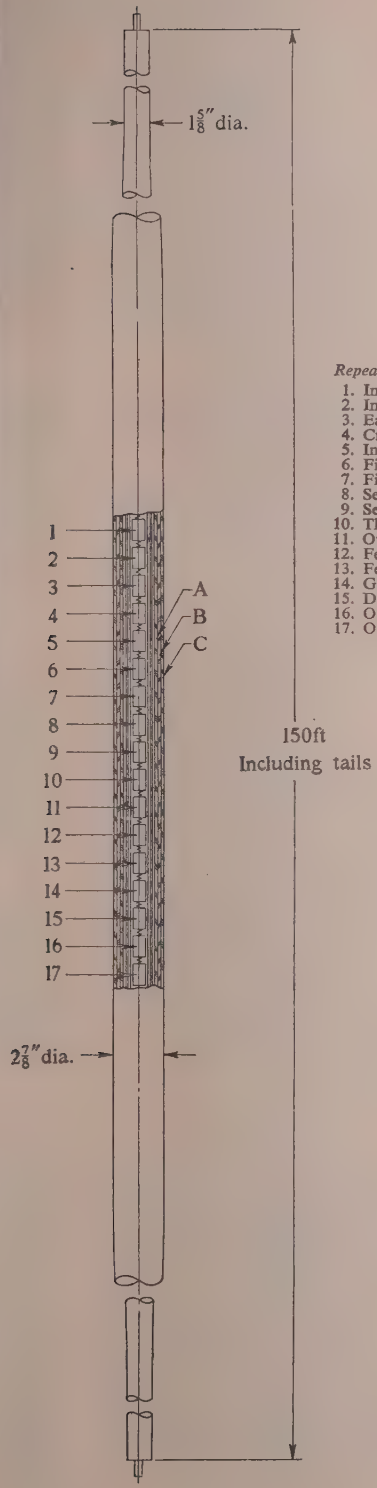
are shown in Fig. 5. Type-A cable has an armouring consisting of 12 galvanized mild-steel wires, each 0·3 in in diameter, and is laid in water not deeper than about 300 fathoms; type AA, used for the landing, is similar, but has an extra layer of armouring. Type B is laid in moderate depths and has 18 galvanized mild-steel wires, each 0·165 in in diameter. Type D is laid in all depths greater than about 700 fathoms; it is armoured with 24 high-tensile-steel wires of 0·086 in diameter, each wire being taped, and these enable it to be laid and lifted in the deepest water.

(8.2) Repeaters

The desirability of a flexible type of housing was referred to in Section 4.1. The structure which was finally evolved is a flexible bulge in the cable 8 ft long and 2⅞ in in diameter, tapering for a distance of about 20 ft at each end down to the cable diameter. The armour of the cable itself is continued over the repeater housing, but with armour wires added to give complete coverage. To prevent the twisting of the repeater structure in the laying operation, there is a second layer of wires with opposite lay. The repeater itself is enclosed within an 8 ft-long series of cylindrical sections inside the housing.

The structure of the repeater container is shown in Figs. 6A and 6B. The whole structure is flexible, and passes readily over the cable drum and bow sheave without requiring that the ship be stopped. To attain this flexibility, the repeater elements within

* The following have been agreed for the overall circuits London–New York or London–Montreal:
(a) Frequency stability better than ±2 c/s.
(b) Noise less than 46 db below 1 mW, read on the C.C.I.F. (1951) psophometer or 38 dba read on Bell System 2B noise-measuring set with F1A weighting.
(c) Minimum crosstalk between any two voice channels not worse than 56 db.
† Slightly displaced to avoid carrier leak.
‡ Cable dimensions refer to the core diameter before the addition of the outer conductor.



Repeater-unit sections.

1. Input terminal.
2. Input blocking capacitor.
3. Earthing capacitor.
4. Crystal.
5. Input network.
6. First-stage pentode.
7. First inter-stage network.
8. Second-stage pentode.
9. Second inter-stage network.
10. Third-stage pentode.
11. Output network.
12. Feedback network.
13. Feedback network.
14. Gas diode.
15. Drier.
16. Output blocking capacitor.
17. Output terminal.

Fig. 6A.—Repeater, Clarenville–Oban section.

A. Armour bedding and corrosion protection.
B. First layer of armour wires.
C. Second layer of armour wires.

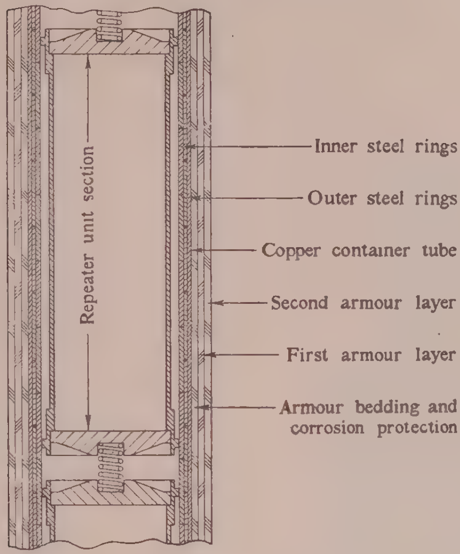


Fig. 6B.—Repeater unit section.

the container are mounted inside a series of Lucite (polymethylmethacrylate) cylinders, successive units being held together by a spring assembly to form an articulated system. Surrounding this series of Lucite cylinders is a series of butt-ended steel rings, in two layers, the joints between successive rings in the two layers being staggered. Over the rings, and supported by them against collapse at sea-bottom pressure, is an envelope in the form of a copper tube, over which are the protective coatings and armouring wires.

The repeater enclosure is terminated at each end with a system of seals, comprising first a glass–metal seal adjacent to the repeater elements, then a plastic seal moulded to the cable insulation and finally, at the extreme ends, a 7ft long seal formed within a copper tube which is an extension of the repeater housing. In order to penetrate the housing, sea water would therefore have to thread a long multi-barriered path. All seals are adapted for the sea-bottom pressures that they may have to withstand. A partially sectioned portion of the seal is shown in Fig. 7, together with further details of the repeater structure.

Fig. 8 is an electrical schematic of the amplifier. The Lucite cylinders into which the electrical elements are packaged are indicated in Fig. 6. Excluding the valves and high-voltage capacitors, 54 components will be accommodated within a volume no greater than that of a 3in cube.

The 52 repeaters in each cable will be located at intervals of approximately 37n.m. They will be fed with d.c. power originating at each shore end, the power feeding arrangements being described in Section 10. Each repeater will produce a voltage drop of 55 volts at a current of 225 mA initially, but after a number of years the current can be increased to a maximum of 245 mA (64 volts per repeater) to offset any deterioration in valve emission.

(8.3) Testing Arrangements

Frequencies in the range 167–174 kc/s will be used to keep track of the performance and condition of individual repeaters by means of measurements made from the shore ends. The techniques for doing this have been outlined in earlier articles describing the Key West–Havana system.[4] Briefly, the device used is a quartz-crystal resonator shunted across the feedback circuit of the repeater so as to remove feedback effectively over a very narrow frequency band which is different for each

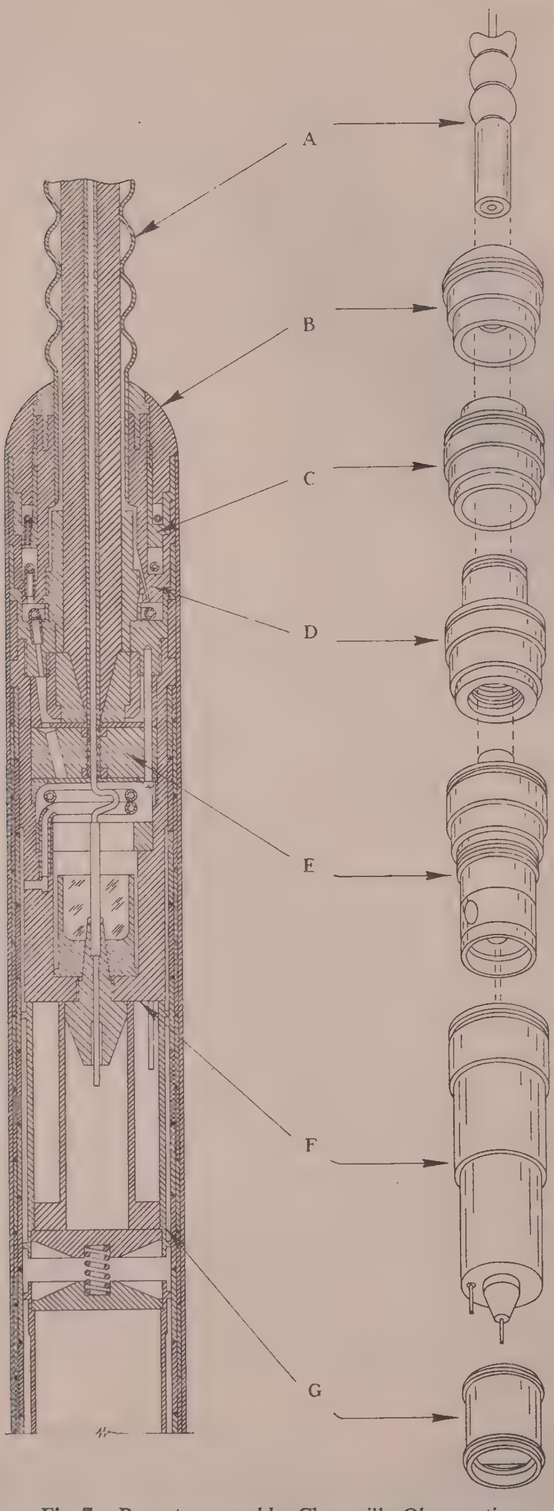


Fig. 7.—Repeater assembly, Clarenville–Oban section.

A. Core tube, 7ft long, Vistac (polyisobutylene) filled and sealed.
B. Nosing parts.
C. Cover.
D. Protective-tube assembly.
E. Rubber-seal assembly.
F. Glass-seal assembly.
G. End-terminal unit.

repeater. In this narrow band the gain is much higher than at all other frequencies and is very susceptible to changes in valve characteristics. Thus, system gain measurements at the various crystal frequencies provide a means for detecting, in each repeater separately, small changes in significant valve parameters which would not be disclosed with normal feedback. At the same time, the high gain at resonance creates a series of thermal-noise peaks at the crystal frequencies. These can be observed by a narrow-band detector at the receiving end of the cable and will provide a means for locating repeaters which have failed. To accomplish this, it is only necessary to note the missing noise peaks, since these are associated with the repeater which has failed and those preceding it. In order to make this test possible even when the failure is caused by an open-circuit heater, a gas diode is provided in each repeater to by-pass the power supply current.

(8.4) Repair Repeaters

If the cable is damaged during laying or at a later time, it will have to be raised and repaired. After the repair, the cable length in the repaired section will be greater than before, and this will upset the balance between cable loss and repeater gain. In the deepest water, the extra cable length may amount to several miles, and to compensate for the additional loss, low-gain "repair repeaters" are being provided. One of these will be inserted when a repair is made in deep water.

(9) CLARENVILLE–SYDNEY MINES SECTION

The frequency bands to be transmitted over the single cable between Newfoundland and Nova Scotia (see Fig. 3) are those used by the British Post Office in most of its shallow-water schemes, except that the bottom frequency band is lowered by 4 kc/s so that the three transatlantic groups can occupy the same frequency band as on the section between Newfoundland and Scotland; this will simplify the frequency-translating equipment.

(9.1) Submarine Cable

The structure and dimensions of the conductors and core of the submarine cable will be identical with those of the cable used between Clarenville and Oban; the cable will have A- or B-type armouring (Fig. 5). It will, however, be tested to a specification appropriate to the higher frequencies to be transmitted and having regard to the fact that some slight easements are possible because the impedance of the repeaters in this section will approximately match that of the cable.

(9.2) Land Cable

The land cable will be basically the same as the submarine cable, but, being more subject to electrical interference, it will be screened with layers of soft-iron tape applied over the outer conductor. The tape in turn will be protected from corrosion by a polyethylene sheath, over which will be applied jute bedding and armour wires. The cable will be laid in a trench at a depth of approximately 30in, where the temperature is expected to have an annual variation of about 35° F and a maximum daily variation of about 2° F. Guard wires, for protection of the system against lightning, will be laid in the trench above the cable. The two repeaters in the land section will be identical with those in the sea and, according to the terrain determined in the final survey, will be laid in ponds, buried in the ground or located in small manholes.

(9.3) Repeaters

The mechanical design of the repeaters which will be used in this section is shown in Fig. 9. The electrical equipment is mounted in an inner sealed cylinder 4ft 2in long and 7$\frac{3}{4}$in

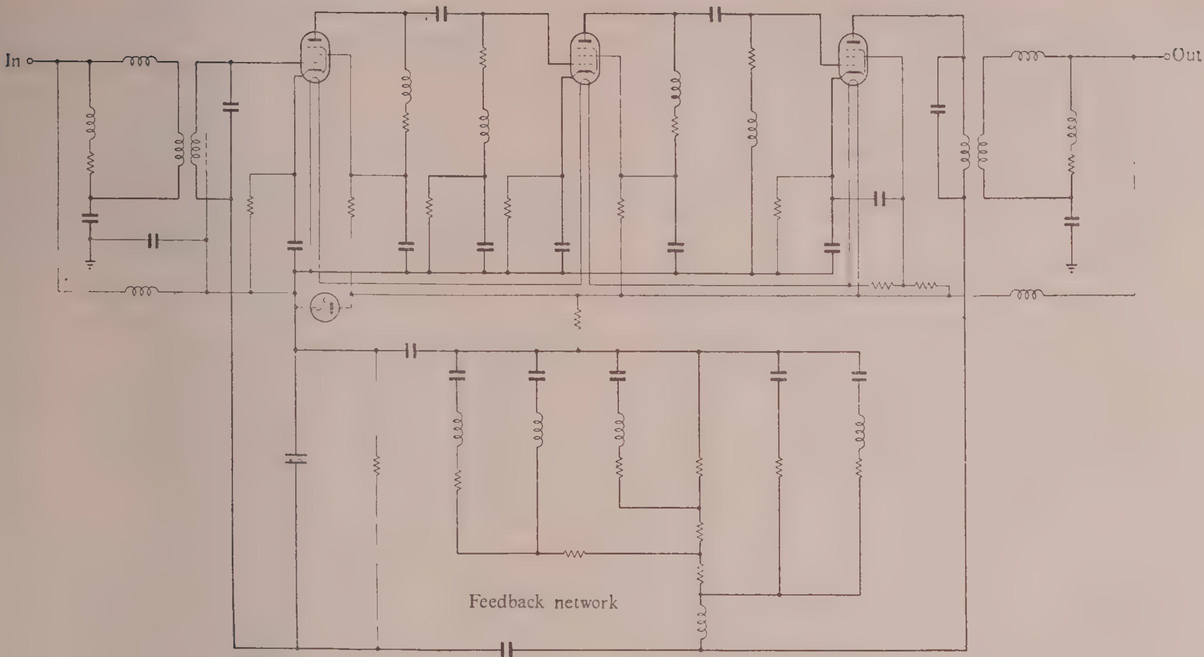


Fig. 8.—Schematic of amplifier for Clarenville–Oban Section.

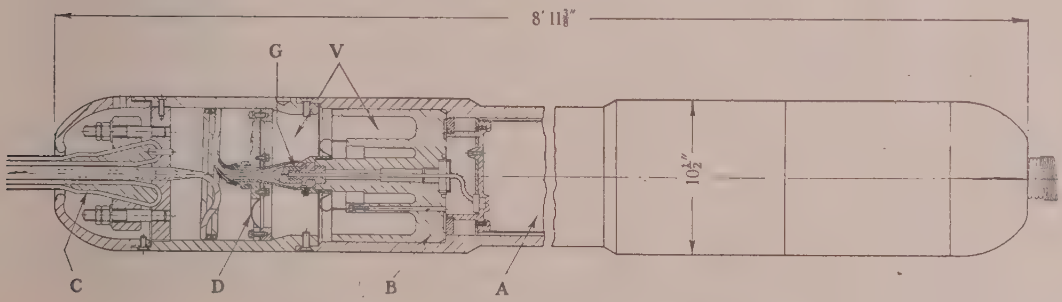


Fig. 9.—Repeater housing for Clarenville–Sydney Mine section.

A. Inner unit enclosing electrical apparatus.
B. Bulkhead (brazed in).
C. Armour clamp.
D. Flexible diaphragm enclosing Vistac, V.
G. Castellated gland.

in diameter, filled with dry nitrogen. This is mounted between brazed-in bulkheads in the pressure-resisting outer housing, which is about 9 ft long and has a maximum diameter of 10½ in. The glands are polyethylene mouldings which are integral with the cable dielectric and which are shrunk on to castellated bosses, previously treated so as to form a bond with the polyethylene. The cable armouring wires are made-off under clamps, which transfer the laying tension to the housing; the design of these clamps is adequate to permit of laying the repeaters at ocean depths, although, of course, such strength will be unnecessary for cable-laying operations between Newfoundland and Nova Scotia.

The 16 repeaters in this section will be located approximately 20 n.m. apart. Electrically, they will be arranged to give both-way amplification, as shown in Fig. 10. A feature of the design is the inclusion of balanced bridges at each end of the repeater, which not only supplement the directional discrimination in the filters by a minimum of 20 db, but also contribute substantially

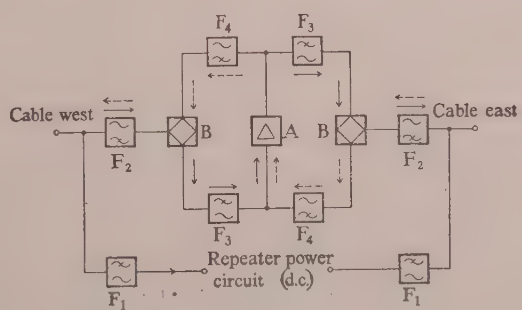


Fig. 10.—Schematic of both-way repeater.

A. Amplifier.
B. Balanced bridge.
F_1, F_2. Low- and high-pass power filters.
F_3, F_4. Low- and high-pass directional filters.
→ West–east channels, l.f. group.
← East–west channels, h.f. group.

to the equalization. The bridges also have the advantage that the intermodulation requirements in the filters are reduced so that small high-Q-factor ferromagnetic-cored coils can be used. Apart from the equalization provided by the directional bridges, the repeater derives most of the necessary gain/frequency characteristic from the two-terminal feedback network.

There are two 3-stage amplifier units in parallel between common input and output transformers with a single feedback path as shown in Fig. 11. The feedback network will maintain

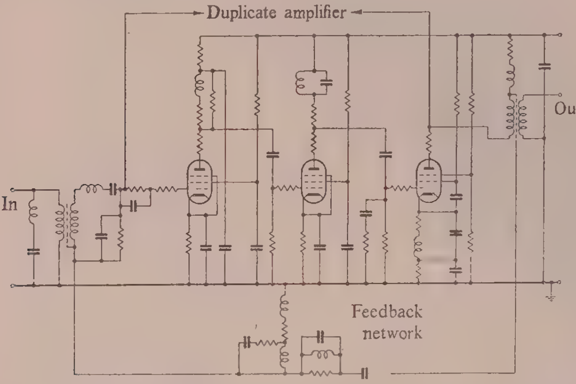


Fig. 11.—Schematic of amplifier for Clarenville–Sydney Mines section.

the repeater gain if one amplifier unit becomes inoperative, but the overload point is reduced by 5 db under this condition; a faulty component in one amplifier unit cannot affect the other unless it introduces circuit noise. In comparison with a system using single amplifiers, the probable life of each repeater is increased and the probability of early faults on the system is greatly reduced, as shown theoretically in the Appendix to an earlier paper.[5]

The wide frequency range to be transmitted over the cable between Clarenville and Sydney Mines will be made practicable only by the use of comparatively modern high-performance valves with a mutual conductance of 6 000 micromhos. It is well known that the mutual conductance of valves using conventional cathode cores—nickel with a small content of silicon and magnesium—decreases with time owing to the growth of a resistive interface between this core and the emissive material of the cathode; this interface applies negative feedback to an extent which increases with age. Work in the Post Office Research Laboratories has shown that the formation of an interface can be prevented by the use of pure platinum cathode cores. Cores of this kind, known as "passive" cores, will be used; these increase the difficulties of processing but result in valves which should have stable characteristics over very long periods of time.[9]

The heater of each valve requires about 300 mA at 5·5 volts, and the six heaters in each repeater are series-connected; the cathode-anode p.d. (90 volts) is derived from a separate resistor carrying the line current. Each group of three heaters is bridged by a device which will short-circuit the group if a heater fails and thus maintain the line current.

The accommodation afforded by the rigid housing not only permits the complicated circuits of the British repeaters, but also makes it possible to design some of the components on more generous lines than can be fitted into the American flexible housings. For example, the power capacitors use a non-polar dielectric of mineral oil instead of polar castor oil and are subject only to about one-third the electric stress. Silvered-mica capacitors will be used only where the polarization is less than 10 volts, and at all other points the capacitors will be of the sealed oil-impregnated-paper type. All components will be manufactured and tested to standards similar to those described in Section 4.3.

(9.4) Testing Arrangements

It is an advantage of a both-way cable that, if some form of frequency translation is possible in each repeater, signals transmitted from one terminal can be returned from the repeaters to the same terminal. In this instance the frequency translation in the repeater is effected, as shown in Fig. 12, by a frequency

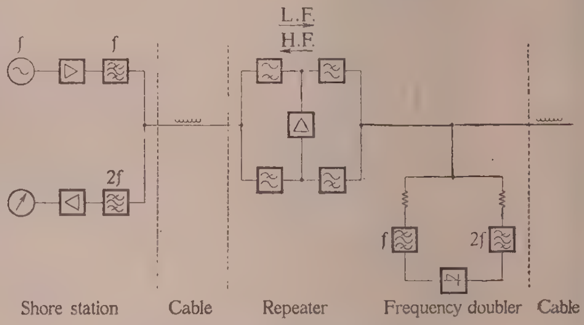


Fig. 12.—Repeater fault location.
Frequency f is in the band 260–264 kc/s.

doubler associated with filters accepting frequencies in the range 260–264 kc/s (i.e. from Sydney Mines) which are individual to the repeater. Signals at twice these frequencies are returned to the sending terminal and enable an accurate measurement to be made of the transmission level at the output of each repeater at any time.

Pulse-testing equipment will also be provided at each terminal to permit measurements of the overload point of each repeater. The repeaters will be identified on a time base and the frequency translation effected by the overload distortion of the amplifier itself. This test can be carried out in either direction, by direct harmonic production in one direction and by intermodulation of a 2-tone signal in the other direction; it is aimed primarily at determining whether both amplifiers are functioning in each repeater.

(10) POWER FEED TO REPEATERS

The arrangements for feeding power to the repeaters on the two sections are generally similar and involve the use of constant-current generators at both terminal stations, series-aiding and arranged so that one end controls the value of the line current (Fig. 13).

The repeaters on the Clarenville–Oban section will require initially a total driving voltage of about 3 900 volts, increasing later to about 4 450 volts, or up to about 4 700 volts if it is necessary to introduce additional repeaters after repairs; the maximum voltage on the cable will be half of this. For the repeaters on the Clarenville–Sydney Mines section the necessary total driving voltage will be about 2 300 volts and, if necessary, it will be possible to feed this from one end only, since the power capacitors are adequately rated.

The constant-current generators used on the two sections will differ in several respects. Those on the Clarenville–Oban section will derive the constant-current feature by feeding via high-impedance pentodes with a mechanical regulator to absorb all but short-term variations. In order to assist continuity of the line current, the load will be shared by two generators at each end, each generator being capable of taking the full load.

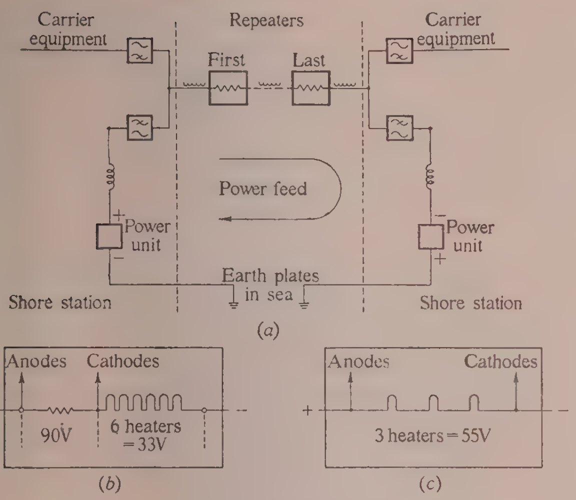


Fig. 13.—Basic power-feeding arrangements.

(a) General schematic.
(b) Repeater power circuits for Clarenville–Sydney Mines section.
(c) Repeater power circuits for Oban–Clarenville section.

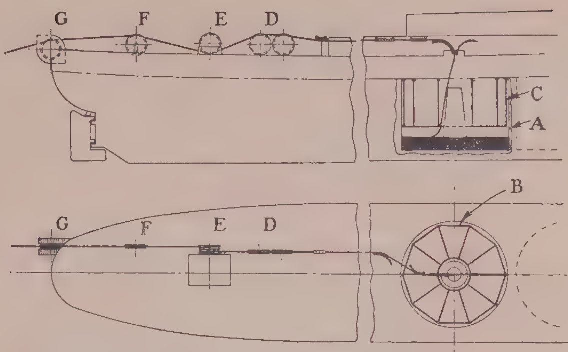


Fig. 14.—Cable-laying machinery on H.M.T.S. *Monarch*.

A. Cable tank (there are four).
B. Crinoline.
C. Crinoline block and tackle.
D. Jockey gear.
E. Paying-out drum.
F. Dynamometer.
G. Stern sheave.

All cable gear has a minimum radius of 3ft 5in.

Transductor-controlled power units will be used for the Clarenville–Sydney Mines section.

Every effort will be made to ensure continuity of the d.c. power supply to the cable, so as to avoid voltage surges at the repeaters, and in order to assist this the a.c. supplies to each power-feed unit will be derived through a.c./d.c./a.c. machines. These machines will be normally driven from the supply mains, but in the event of mains failure, Diesel-engine-driven alternators will be started up automatically. There will be two stages of standby:

(a) The a.c./d.c./a.c. machines will continue to run, driven from a 130-volt battery which energizes the d.c. machine.
(b) If the mains failure is of long duration, the standby Diesel plant will take over.

(11) CABLE-LAYING ARRANGEMENTS

All the cable will be laid by H.M.T.S. *Monarch*. This ship, which was built for the British Post Office in 1945,[10] has a gross tonnage of 8 056 and, with full oil bunkers, can carry between 5 000 and 6 000 tons of cable.*

Monarch is the only cable ship afloat capable of laying the whole of the deep-water part (about 1 500 n.m.) of each cable between Newfoundland and Scotland in one operation. From earlier Sections of the paper it will be obvious that it is desirable to avoid any sequence of operations which will make it necessary to pick up a cable end and to make a splice in deep water. Such operations will have to be carried out if repairs are necessary after the cable has been laid, but inevitably introduce some risk of the cable kinking; the initial laying will be planned to avoid them. All repeaters will have been jointed into the cable before it is loaded into the ship.

A period of 12 days during which continuously good weather conditions may be anticipated is necessary in order to carry out the main operation without undue hazard to the repeatered cable. In the North Atlantic such conditions are unlikely, except during the period from mid-May to mid-September, and it will not be practicable to lay both the cables on the main crossing in one summer. It is planned to lay one of these cables during the summer of 1955 and to complete all cable-laying operations in 1956.

* The tons referred to here, and in the paper generally, are long tons of 2 240 lb.

The essential parts of the after cable paying-out machinery on *Monarch* are shown in Fig. 14. Although the flexible built-in repeaters are designed to be handled as part of the cable, it is, of course, desirable not to bend them to a small radius. New cable drums of approximately 7 ft diameter are being fitted on the *Monarch*, this being the largest size that can be accommodated without major structural alterations. Bow and stern cable sheaves will also be made this diameter. Roller guides with a minimum radius of 3 ft 6 in will replace the existing spider wheels and "spectacle" guides throughout the ship, in order to ensure freedom from sharp bends and an easy run of cable from the tanks during laying. Various devices which proved successful when laying the Key West–Havana cables will be fitted to ensure that the repeaters conform to the curvature of the cable drums and to ease their passage round the drums and at other points.

Three new dynamometers will also be fitted, two forward and one associated with the paying-out gear aft. Sheaves of 7 ft diameter will be used, associated with electrically-sensitive pressure cells. The latter will be arranged to operate indicating and recording instruments on deck and in the test room, and will enable the tension on the cable to be controlled closely as it is paid out.

The laying of the cable between Newfoundland and Nova Scotia will present no special problems, but an electrically driven hoist has been fitted over the ship's bows to facilitate handling the rigid repeaters.

(12) CONNECTION OF CABLE SYSTEM TO AMERICAN, CANADIAN AND BRITISH (AND EUROPEAN) INTERNAL TELEPHONE NETWORKS

A description of the cable and radio-relay systems which will be used to extend the transatlantic circuits from Oban to London and from Sydney Mines to New York and Montreal would be outside the scope of the paper, and in any case, systems of the same kind have already been extensively used in the countries concerned. It may be of interest, however, to indicate how the connections will be made. From Oban the transatlantic circuits will be taken to London by carrier cable; there will be two alternative routes, one via Glasgow and the other via Inverness and Aberdeen. In London it will be possible to connect any transatlantic circuit to any one of the existing submarine-cable circuits to the continent of Europe, thus providing through cable communication between North America and continental Europe.

On the American side a microwave radio-relay system will connect Sydney Mines with Portland (Maine). This type of system, which makes use of line-of-sight transmission between relay stations, has already been extensively installed in the United States as a means of providing long-distance telephone circuits of the same quality as cable circuits. From Portland the circuits will be extended to New York by carrier cable or radio relay. Cable connection to Montreal will be effected at a point on the radio-relay system near the United States–Canada border.

(13) CONCLUSION AND SPECULATION

Barring unforeseen and unexpected delays, 1956 should see the completion of the first transatlantic telephone connection via submarine cable. With the completion of this project, telephone service between the two continents should enter into a new era marked by improved quality of service and reliability, and a capacity for growth no longer restricted by the limited capacity of the radio spectrum. However, the advantages of radio—flexibility, the speed with which communications can be established and shifted from route to route, the feature of direct access from one country to another, relatively low cost, etc.—will remain. These are not inconsiderable advantages and they augur well for the continued importance of short-wave radiotelephone services.

On the other hand, the new transatlantic cable will by no means represent all that could be done if full advantage were taken of contemporary techniques. The submarine-cable art is a conservative one. As Dr. Buckley observed, the laying of a submarine cable, like any other activity on the high seas, must be planned with its unavoidable hazards in mind. Experience over 100 years has shown that when handling deep-water submarine cables on the high seas it is better to be safe than sorry.

But the consequences of this doctrine of caution are equally plain. The submerged repeaters in the 1956 deep-water-cable section incorporate valves whose proven design dates back to 1941, and with a mutual conductance of 1 000 micromhos. On the other hand, "conservative" post-war valves provide mutual conductances of about 5 000 micromhos, and valves having a slightly better performance than this will be used in the cable system between Newfoundland and Nova Scotia under conditions where the consequences of valve failure will be minimized. The integrity of such post-war valves under the severe and highly specialized conditions of installation and use encountered in submerged repeaters remains to be established. On the other hand, it is without question that, with such new valves, cable systems having still greater communication capacity can be realized.

Looking still further ahead, the transistor looms as a development which has the potentialities for making possible long deep-water submarine cables with much greater communication capacity than can be realized with repeaters employing thermionic valves. The voltage required by the valves sets an upper limit on the number of them that can be operated in tandem before the power supply results in voltages on the cable that are beyond the limits of safety. The much lower power drain of the transistor should overcome this obstacle and permit more repeaters to be used and wider frequency bands to be transmitted. Ruggedness, long life and small size are added attractive features of the transistor which are of not inconsiderable importance to the future development of submarine cables.

A transatlantic submarine television cable is a long-range goal worthy of serious study and by no means to be dismissed as impractical of eventual attainment.

(14) ACKNOWLEDGMENTS

The concept of a transatlantic telephone cable owes its origin to the inspiration of Dr. Buckley's early work, for he laid the firm technical foundation for such a project. The subsequent developments recorded in the paper have been the work of many engineers on both sides of the Atlantic; on the American side in the Bell Telephone Laboratories, the Western Electric Company and the Simplex Wire and Cable Company; on the British side in the Post Office Research Laboratories, Standard Telephones and Cables, Ltd., and Submarine Cables, Ltd. Without this work the particular project described would have been impossible.

(15) BIBLIOGRAPHY

(1) Buckley, O. E.: "The Future of Trans-Oceanic Telephony," *Journal I.E.E.*, 1942, **89**, Part I, p. 545.

(2) Garratt, G. R. M.: "One Hundred Years of Submarine Cables" (H.M. Stationery Office, London, 1950).

(3) Radley, W. G.: "Fifty Years' Development in Telephone and Telegraph Transmission in relation to the Work of Heaviside," *Heaviside Centenary Volume* (Institution of Electrical Engineers, 1950), p. 76.

(4) Gilbert, J. J.: "A Submarine Telephone Cable with Submerged Repeaters," *Bell System Technical Journal*, 1951, **30**, p. 65.

(5) Halsey, R. J., and Wright, F. C.: "Submerged Telephone Repeaters for Shallow Waters," *Proceedings I.E.E.*, Paper No. 1633, February, 1954 (**101**, Part I, p. 167).

(6) Walker, D. C., and Thomas, J. F. P.: "The British Post Office Standard Submerged Repeater System for Shallow Water Cables with Special Mention of the England–Netherlands System," *Proceedings I.E.E.*, Paper No. 1634, February, 1954 (*ibid.*, p. 190.)

(7) Roche, A. H., and Roe, F. O.: "The Netherlands–Denmark Submerged Repeater System," *Proceedings I.E.E.*, Paper No. 1635, February, 1954 (*ibid.*, p. 180).

(8) Halsey, R. J.: "Modern Submarine Cable Telephony and the use of Submerged Repeaters," *Journal I.E.E.*, 1944, **91**, Part III, p. 218.

(9) Metson, G. H., Wagener, S., Holmes, M. F., and Child, M. R.: "The Life of Oxide Cathodes in Modern Receiving Valves," *Proceedings I.E.E.*, Paper No. 1137 R, June, 1951 (**99**, Part III, p. 69).

(10) Gill, A. J.: "H.M.T.S. *Monarch*," *Post Office Electrical Engineers' Journal*, 1947, **39**, p. 129.

DISCUSSION BEFORE THE INSTITUTION, 4TH NOVEMBER, 1954

Dr. O. E. Buckley (*United States: recorded*): The system described represents the most radical and important advance yet undertaken in the old and conservative art of transoceanic cables. One way to measure this advance is by the top frequency which can be transmitted. It took more than half a century to go from the 1 c/s or so of the first transatlantic cable to the 100 c/s or so of the continuously-loaded telegraph cables. Now it is planned to go in one great step to a top frequency exceeding 100 kc/s and to transmit a broad band of frequencies to provide 35 or more telephone circuits.

One reason for the traditional conservatism of the cable art is the great cost in time and money of making repairs once a cable is laid. This has put a high premium on ruggedness and simplicity. In this new project, simplicity will be sacrificed to no small degree by the incorporation in the cable structure of a complicated assembly of electronic equipment precisely designed to compensate for the characteristics of the cable over a wide range of frequencies; and there will be not just one such repeater, but over 100 of them in the two cables for the Newfoundland–Scotland link. This calls for unprecedented care and precaution

to ensure success. It has called for extraordinary devices and for imagining all possible causes of trouble that may show up and for having means to deal with them.

One method of making such a repeatered cable was proposed in the 1942 Kelvin Lecture,* and repeaters of the type described therein, with important improvements subsequently added, will be used in the long deep-sea section of the system. The link to connect Newfoundland and Nova Scotia will employ two-way repeaters developed by the British Post Office. Experience to date with installations of both types of repeater has demonstrated their practicability.

So radical a venture can be justified only by great value to be gained, and that value is indeed very great in the present instance. Although the transatlantic telephone service now provided by short-wave radiotelephone systems has become commonplace, it is still short of the perfection of long-distance overland telephone transmission provided by coaxial cable and radio relay systems. Furthermore, oversea short-wave radio is on occasion subject to impairment or complete interruption by magnetic storms. Even when radio circuits are working well, there are not enough channels available in the short-wave spectrum to satisfy the ever-growing demand for public service. The provision of this telephone cable service across the Atlantic will surely lead to greater demand; this, in turn, will lead to more cables, and I expect that cables will become the backbone of transatlantic telephone communication.

Confident as I am in the practicability of repeatered submarine cables and the need for them, I am equally sure that the present designs are not the ultimate and that still better cables will follow. Just as this first edition of a broad-band transoceanic telephone cable system has been made possible by new materials and techniques hatched in the laboratory, so will improvements, many of which are now under way, come into being.

Making the present project a joint undertaking brings to it the products of two similar but different lines of development. This is significant of the spirit which has long characterized the co-operation between the Bell System engineers and the engineers of the British Post Office. Out of this co-operation will come benefits which will contribute to future progress.

Col. Sir A. Stanley Angwin: Having been particularly interested in the first development of transatlantic telephony by radio and in the first proposals for a submarine telephone cable, I am very pleased indeed to have the opportunity of contributing to the discussion on the project outlined in the paper.

This paper, with its picture of the growth of the North Atlantic wireless system and account of the first work on submarine repeaters, is of particular interest to The Institution as a corollary to the 33rd Kelvin Lecture† given by Dr. Buckley in 1942, which I had the pleasure of presenting; and I myself had the opportunity, in my Presidential Address‡ in 1943, of giving some indication of what was being done and contemplated in the United Kingdom. (That Address was, of course, made with many restrictions as to what could be published while the war was in progress.)

The tremendous impetus given to long-distance wireless from 1928 onwards stimulated the cable interests into consideration of the possibility of long-distance submarine cable telephony. An indication of the progress made is the extension of the bandwidth from that of the first transatlantic cable, 1·5c/s, to 100 c/s on the continuously-loaded telegraph cables, and to 144kc/s on the new telephone cable.

The conclusion has now been reached that the ideal should be the judicious combination of radio and cable for transoceanic conditions. In this particular case, not only will telephony by cable be added to radiotelephony between the United Kingdom, Canada and the United States, giving a much more efficient service at all times, but the possibility is added of telephony from the United Kingdom to Australia and New Zealand over the cable and the trans-Canada line network, and thence by radiotelephony on the Pacific route, thus improving the through telephony service to the antipodes to the extent that it will operate over those periods when it is impossible on the direct radio link.

One of the outstanding factors that may be deduced from the paper is that the mechanical problems which have to be met in this undertaking are more critical than the electrical ones. Fundamentally, the solution of the repeater is similar in the United Kingdom and in the American design. The greater difference is in the housing—one articulated and flexible, the other bulkier and rigid. There is a feeling of confidence in the design on the electrical side, a confidence inspired by exhaustive research and experiment without which this great work could not have been undertaken. On the mechanical side the conditions to be met are less determinate and the risks difficult to cover by a factor of safety. This is perhaps the most difficult feature and one to which much thought and trial have been given.

Other problems are more intractable. The difficulties of cable laying in the Atlantic at depths down to 2 300 fathoms are touched on in the paper. There have in the past been many epics of cable laying and cable repairs, from those of the laying of the first transatlantic telegraph cables to the more recent attempt to restore the Porthcurno–Harbour Grace cable in 1944, when 36 "drives" were made by the cable ship in a period of 40 days in conditions of fog and storm and other difficulties which prevented successful repairs. Indeed, restoration was not finally carried out until 1952. I feel that, when the first transatlantic telephone cable has been well and truly laid, great credit will be due to the seamanship of those who overcome these hazards of nature. I hope that this phase of the undertaking will be recorded and in due course brought before The Institution.

Sir Archibald Gill: Sir Stanley Angwin has called attention to the fact that many of the problems involved in this project are mechanical. One of the greatest problems, of course, concerns the gland in the repeater, and I feel that more detail might be given of that very interesting item. I have understood most of the paper except the reference to the material Vistac.

Sir Stanley has referred to the fact that this is a joint paper, submitted to The Institution and to the American Institute. Many will recall that, shortly after the first transatlantic radiotelephone service was opened in 1927, we held a joint meeting of The Institution and the American Institute of Electrical Engineers which was conducted over that radiotelephone service. I should like to suggest that, when this project is completed, another joint meeting between the two bodies should be held. I feel certain that both the British Post Office and the American Telephone and Telegraph Company will consider the suggestion sympathetically.

I have been interested in some of the financial implications of this project. Looking at the length of cable involved in the transatlantic section and knowing little about the present cost of these things, I should guess—it is only a guess—that the cost of this part of the project will not be much less than £7 000 000, and it may be more. Taking a figure of £7 000 000 and allowing interest at 3½%, we get a figure of £250 000 per annum for interest, while amortization over 35 years will mean another £200 000 per annum, or a total of £450 000. To maintain the cable it will probably be necessary to have a special ship costing something of the order of £100 000 per annum, thus making £550 000 per annum.

* BUCKLEY, O. E.: "The Future of Transoceanic Telephony," *Journal I.E.E.*, 1942, 89, Part I, p. 454.
† *Op. cit.*
‡ *Journal I.E.E.*, 1944, 91, Part I, p. 15.

Then there are the operating charges. I do not know very much about that, because it is not engineering, but I believe that, to get full value from a cable of this kind handling personal calls, it is necessary to have two operators at each end, because while one operator is connecting the subscriber and monitoring the call the other is making arrangements over the circuit for the next call. This means four operators for each of the 36 circuits during the busy hours, although fewer would be needed at other times, say a total of 200–250. I cannot see how they can be provided, including overheads, at a cost of less than £120 000 per annum, which brings the total to £670 000 per annum.

There are only 250 business days in the year, excluding Saturdays, Sundays and public holidays, so that to cover the cost it would be necessary to take about £2 700 per day. There are only 3½ hours in the day, because when it is 9 a.m. in New York it is 2 p.m. in London, and when it is 1 p.m. there it is 6 p.m. here, and the New Yorker is going to lunch and the Englishman is going home for the day. The hours available for calls, therefore, are really from 2 p.m. to 6 p.m. G.M.T.; allowing a quarter of an hour at each end to warm up gives about 3½ hours' operation each day. Even with two operators per circuit per end, it is not practicable to get more than 30 paid minutes out of an hour on circuits such as these, so that the total paid minutes per circuit per day would be about 100.

The present charge for a call from Britain to Canada by radio is £1 per minute, and at this rate each circuit could earn £100 per day during the business hours—a total of £3 600 per day—so that costs would be covered as soon as that amount of traffic was reached. When the cable is first opened, of course, there will not be enough traffic to occupy all that time, but as soon as the traffic approaches about 75% of the cable capacity the cost ought to be covered. During week-ends, evenings and nights there will be a certain amount of traffic which I have not included, which will increase the income. On the other hand, there will be additional costs for the links from Newfoundland to the mainland of Canada and onward to New York. On the whole it looks as if the cable will be able to pay its way.

Some reference has been made in the paper to cable ships, and particularly to H.M.T.S. *Monarch*. The *Monarch* was built during the last war, when we could not try very many experiments, so that she is similar generally to the ships which preceded her. If a new cable-laying ship were to be built, I think that there are other improvements which could be included. One possibility would be to arrange the ship rather like an aircraft carrier, with the funnels and bridge structure on the starboard side, leaving the deck completely free on the port side. This would give ample deck space for handling and jointing and make it possibly easier to transfer cable brought in over the bow sheaves to the stern sheaves. In many operations, e.g. landing a shore end, it is more convenient to pay out over the bow sheaves, but later when laying in deeper water it is preferable to pay out from the stern. At the present time, to transfer a cable from the bow to the stern it is necessary to cut it, pull it round the ship and joint it again at the stern—a process taking several precious hours. Another feature worth considering is the adoption of a wide flared bow like that of an aircraft carrier, making it possible to keep the sheaves well away from the centre-line of the ship, which, I think, would be of advantage.

Mr. C. C. Duncan (*United States*): It is a pleasure to be here on this occasion with some of my colleagues in the Bell System who are working on this project. There are many others in the Bell System who would like to be here, such as Dr. Kelly and Mr. Gilman, men who are greatly interested in and actively engaged on the scheme. In fact, the public in our country have had their imagination stirred by the project, and we sometimes feel that we have a million eyes looking at the back of our necks. However, with the magnificent co-operation which everyone is giving to the scheme, including the British Post Office, the Canadian Overseas Telecommunication Corporation, the Bell System, and the manufacturers of the cable and equipment on both sides of the Atlantic, and with the ability of the people assigned to the scheme by these organizations and their intense desire to get the job done successfully and on time, I am fully confident that its challenge will be met.

Mr. A. C. Hartley: The mechanical engineering proposal for pumping petrol across the English Channel during the war was carried through to success as Operation Pluto* only because of the great help given throughout by the Post Office Engineering Department and by the cable industry. The electrical engineering project described in the paper is one in which the inevitable hazards and risk of damage are greatly increased by any lack of continuity in the lay. My Pluto experience makes me very conscious of the great need for continuity.

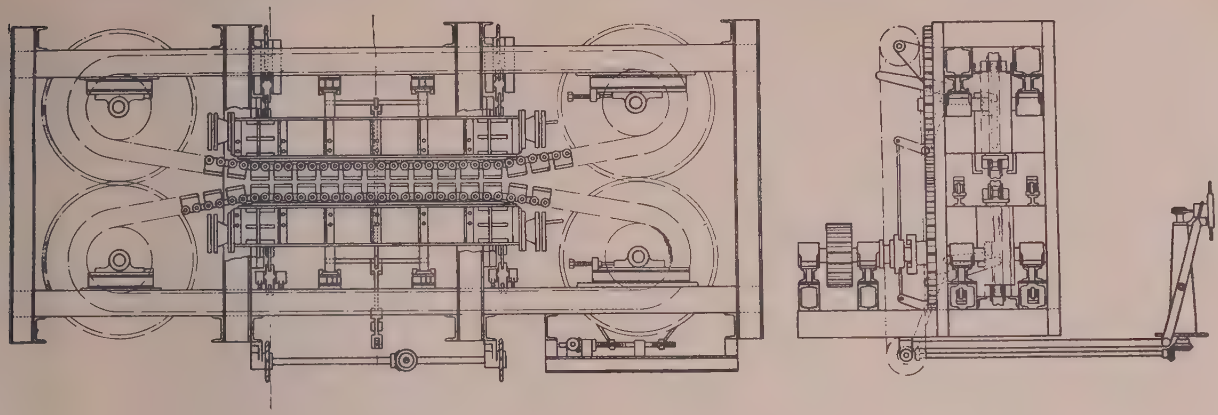
The U.S. snake-type repeater is a most ingenious step towards achieving this continuity of lay, but it is stated in Section 6 of the paper that "it was apparent that, if means for laying and recovering the large repeater housings in deep water without damage to the cable were later developed, the design"—that is, the British development—"would provide great flexibility in repeatered cable systems of the future." The mechanical engineer will certainly take up this challenge and will not be satisfied until he has devised means for the continuous handling, from the cable tank into the sea, of the type of repeater best suited for the electrical engineer's requirements.

Various proposals for achieving this continuity of lay are being considered. Fig. A shows a caterpillar design of which a simple version was used for Pluto. The cable is controlled by grips fixed to the links of two endless chains which can be driven or braked. There is hydraulic control of the pressure of the grips on the cable. When the repeater is to be passed through, the hydraulic pressure is released, wedges are withdrawn and the grips can be opened to allow the repeater to go through. Two complete units would be used in tandem, one controlling the cable and the other open to allow the repeater to be passed.

There is, however, a great attraction in simple rotary motion, and Fig. B shows the first stage in an attempt to use cable drums or sheaves. It is impossible to take the rigid repeater round the cable drum, so why not take the cable drum round the repeater? The first thought was to mount two drums, each on two axles carried on a frame, which would enable the whole to be rotated about a centre between the two axles. The two inner drums were geared together, so that they would rotate in the same direction, and the two outer drums were free to rotate on their axles. Fig. B(i) shows the repeater being passed between the drums, after which the handle to the right of the illustration would be rotated and the cable astern of the drums would be picked up by the two inner drums, and the cable forward of the repeater would be picked up similarly by the outer pair of drums. Several turns of cable can be put on both drums in this way, as shown in Fig. B(ii).

However it was soon proved by calculation that if the drums are suitably grooved it will be possible to get full control of the cable without a complete rotation of the system. Experiments with 6 ft-diameter sheaves geared together to rotate in opposite directions demonstrated the accuracy of these calculations. It was therefore possible to dispense with the four drums and to use two simple sheaves, as shown in Fig. C(i). The two forward sheaves at the right-hand side of the illustration are one above the other, providing room for the repeater to pass between

* Hartley, A. C.: "Operation Pluto," *Proceedings of The Institution of Mechanical Engineers*, 1946, **154**, p. 433.



(i) (ii)

Fig. A.—Caterpillar winch for paying out submarine cable.

(i) Side elevation.
ii) End elevation.

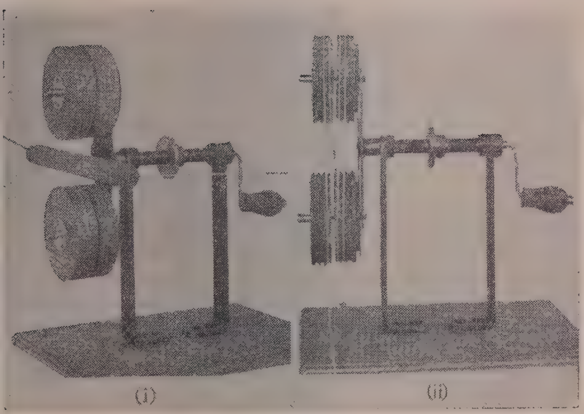


Fig. B.—Model of paying-out drums.

(i) Repeater passing between drums.
(ii) Turns of cable on drums.

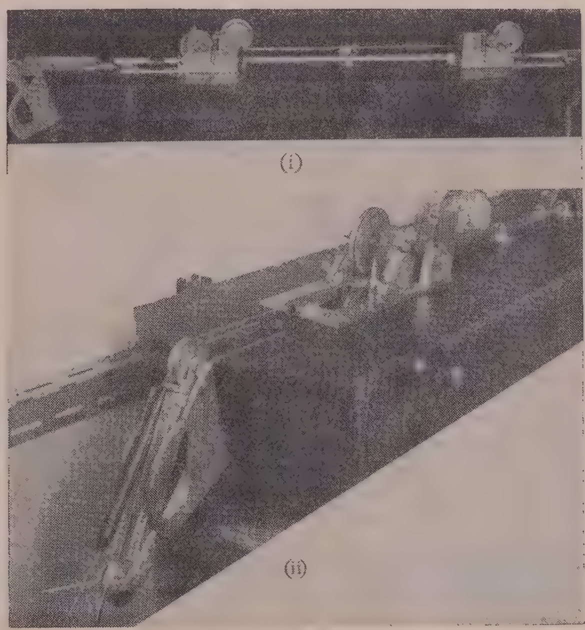


Fig. C.—Model of paying-out gear.

(i) Repeater about to enter forward sheaves.
(ii) Repeater about to be launched.

em, while the two aft sheaves are in the erect position, providing ntrol of the cable. The frame carrying the sheaves can be tated by the worm-and-worm gear shown in Fig. C(ii), about o centres, one of which is midway between the sheaves and the her on the axle carrying the upper sheave. In Fig. C(ii), the sheaves have been opened to allow the passage of the repeater d the repeater is seen about to be launched into the sea from e special launching gear. In order to prepare for the passage the next repeater, the launching gear will be hauled back to e position shown in Fig. C(i) and the aft sheaves will be rotated that the axle of the upper sheave engages in the bearing; the ame carrying both sheaves is then rotated about this centre, inging the lower sheave in a clockwise direction to pick up d control the cable, when it will be in the position shown in g. C(i). The change of the centre of rotation and the launching ar enable the repeater to be handled solely by a straight pull the cable near the repeater glands, and the repeater travels a straight line until it enters the launching gear.

The authors have said, and Dr. Buckley has agreed, that the bmarine-cable art is a conservative one, and that experience er a hundred years has shown that when handling deep-sea submarine cables on the high seas it is better to be safe than sorry. Mechanical engineers are determined to develop gear so reliable and so easy to operate and maintain that it will win the confidence of marine engineers to such an extent that, after one or two lays, they will no more think of dispensing with this type of laying gear than they would think of dispensing with steam propulsion and going back to sail. Electrical engineers may therefore confidently design the repeaters they want and they will be handled without shock.

Major-General Sir Leslie Nicholls: Why is the route for this cable entirely undersea? I am sure there must be a very good reason for this, but could unattended shore repeaters have been

chosen and the cable landed in Iceland and Greenland? Sir Archibald Gill quotes the cost as £7 000 000, but I believe that the estimated cost is £12 500 000, so that his arithmetic will have to be revised.

The use of the one-way snake-type repeaters, as opposed to the British pot-type of 2-way repeater, means that the cable has to be duplicated, which must increase the cost of the project enormously. We in Cable and Wireless are naturally very interested in it, because we feel sure that the project will be successful and that it is going to set a pattern for repeater telephone cables, e.g. to South Africa and South America via Gibraltar or Bathurst. It seems that some three or four years must elapse before another project of this magnitude can be engineered. However, when we have more information as to its success or weaknesses, it will set the pattern, and I hope that this country will be the first in the field with telephone cables to the Commonwealth. There should be an enormous future for them.

Brigadier L. H. Harris: Several speakers have referred to mechanical engineering difficulties. We did have difficulties with glands and casings, but now we feel that we have solved the gland problem once and for all and we are looking forward to having machinery that will stream out rigid repeaters much as Mr. Hartley describes.

The worrying problem from the earliest days of shallow-water repeaters has been to find components and valves with a predictable life of 20 years. Even in shallow water the loss per fault is of the order of £5 000, and bearing this in mind I remember asking the managing director of one firm the price of a capacitor with a £5 000 guaranteed life of 20 years. He said that it would be £5 000. Two years ago, when the Post Office and the Bell System had to decide upon the engineering of the system described in the paper, the latter was in a position to stake its reputation on 6 000 components. In the United Kingdom we were catching up rapidly, particularly with our valve, but had not the years of evidence behind us. It was not difficult, therefore, for the Post Office engineers to accept this already proven system, although it necessitated two cables in the main span against our own preference for a system of two cables each with both-way repeaters. Although we are satisfied with our components now, it would be useful if some neutral reference laboratory had the task of continuous life-testing of all components as they came on the market.

The number of circuits to be obtained from a system is proportional to the square of the number of repeaters, and this "carrot" may lead to over-ambitious systems. The determining factor must be the minimum repeater spacing (and this I believe should be more than that given by simple proportionality with the depth of water) which will give prospect of maintaining the system without appreciable accumulative degradation over the years. This must be so until such time as we have a cable with distributed amplification or some system of automatic regulation to avoid the problem of maintaining meticulously spacings and levels in the face of the whims of the unruly elements which oppose our cable ships.

Mr. J. N. Dean (*read by Dr. E. W. Smith*): My experience as a maker of submarine cables leads me to believe that in this art the electrical engineer must lean towards chemistry and physics, and certainly the chemist must have strong leanings towards electrical engineering, although I do not want to suggest that either of them should lean too much or too often. I believe, however, that there is a tendency to treat this subject too much from the electrical standpoint and not enough from the chemical and physical ones. A project of the nature of the transatlantic cable, apart from the vital importance of its economics, national prestige and safety in defence, involves pure science, particularly chemistry and physics, and also chemical engineering, to a ve great extent.

The teams on both sides of the Atlantic have done a wonder job in combining all the sciences to make this project one of t outstanding advances in communications of our time. To this they have included an intense study, down to the minut detail, of the chemistry and physics of all the components of t cable and equipment, and also of the mechanical and chemi engineering involved in their manufacture. In achieving t they have both drawn on the valuable work contributed other equipment manufacturers and the cable makers. Mo over, what must not be forgotten is that they have drawn heav on the traditional knowledge built up by the pioneers who la the early cables and of the engineers, naval architects and s captains who have kept the knowledge up to date.

Dr. E. W. Smith: It will be noticed that the centre conduct of the cable is a composite structure using three tapes. Brit practice for some time has been to use six tapes, possibly becau of the imagined necessity for more mechanical security, b also because six tapes were easier, in the British view, to fo round a wire with the necessary precision. While I have quarrel whatever with the present design of the cable a thoroughly support the continuance of a type which has been well proved over so many hundreds of miles, I feel that in t future there is a strong case for replacing this composite co struction by a single solid conductor.

This type has already been adequately proved by its use 2 000 miles of really deep-sea telegraph cable, and has actua been used on some shallow-water coaxial cables. One of advantages is that it saves at once about 3% of the attenuati of the cable. I have not found any real evidence of bet mechanical protection in conductors by the composite structu and I do not know of any case of continuity being maintai through one or more unbroken tapes in an otherwise fractu conductor. I except certain joints of early type, now obsolet

In the repeater there is a 7 ft tube through which the ca passes to the interior, and which is filled with Vistac. T prevents the ingress of water from the outside, but if the co ductor still continues down this tube as a composite structur think that it will allow water to penetrate almost as far as glass-metal seal. This type of conductor, if tested with open end in a cylinder under hydraulic pressure, will pass wa at the rate of several hundred cubic centimetres per hour; a since the water will arrive eventually in the conductor diffusion through the polythene, it seems that it will get v close to the glass-metal seal. This component will no dou look after the danger, if it exists, but this is another argum for the solid conductor, which contains no crevices to fill w water.

In the flexible repeater the space available for the electri circuits seems to be about 200 in^3, while in the rigid type i no less than 2 000 in^3. I appreciate that much of this differe is due to the 2-way circuit in the latter case, but how much actually accounted for by the directional filters and so fo and how much is the result of the designers' feeling that, f from the necessity to make the repeater behave like cable, t can give themselves real spatial comfort?

The authors say that the first coaxial cable of modern t generally similar to that described was laid in 1937. This insulated with paragutta, and was therefore sufficiently differ from the present type to excuse a further retrospect to 19 when coaxial cables were laid across the Straits of Gibraltar a between the islands of Tenerife and Gran Canaria. Althou provided with an outer conductor of a large number of wi these cables were characterized by a notable increase in dielectric thickness compared with past practice. The ratio

diameter of conductor to that of insulation was 3·25, which is not too far removed from the calculated optimum—from the attenuation point of view—of between 3·5 and 3·6.

Captain W. H. Leech: We hope that this transatlantic cable will be working at the end of 1956, and every care has been taken that the route selected should avoid known trawling grounds, ships' anchorages and the possibility of damage by grounding icebergs. Whilst we do not expect any faults, we must have regard to the fact that we may have a catastrophic failure, and that we must be able to repair these cables as quickly as possible should such a thing happen. Apart from the *Monarch*, there is only one ship in the vicinity of the North Atlantic capable of repairing either of these two cables during the winter months, and I think that we therefore need one additional cable repair ship, if not two, able to go out into the Atlantic to effect a repair in all weathers. I do not quite agree with Sir Archibald Gill when he suggests a "flat top" for a cable ship; I prefer masts and funnel on the centre-line.

We need a cable ship of about 4 000 tons, with a speed of 15 knots. A speed of 10–11 knots is too slow when it may be necessary to go 1 000 miles to reach a fault. Provision for all our equipment, including larger cable buoys, heavier cable-laying machinery, testing-room space and possibly machine shops in which repairs to the repeaters can be made will be necessary, and we also want the ship to be able to stay in the Atlantic in the winter for long periods to complete the repair.

I know that it is far simpler if we can have machinery which will pay out a cable with rigid repeaters without stopping, and I am very happy to learn that the mechanical engineers are solving that problem. Our present gear for paying out aft is very simple; it is really a drum with a brake. Our equipment has to operate under the most arduous conditions: it must work in the tropics, when possibly the bitumen of the cable is running off over the working parts of the gear, and it has to operate in northern latitudes, in snow and icing conditions; it must also run without stopping, because we cannot stop to repair it. I therefore ask the mechanical engineers to make any equipment simple, robust, easy to maintain and very reliable. Before we install it on a ship it should be thoroughly tried out ashore under all the conditions which we meet with at sea, with salt water playing on it at varying temperatures, so that we can be sure of getting something very reliable. The seven repeaters which we laid recently in the North Sea were laid in boisterous weather—almost a gale—and at night as well as during the day. We are confident that we could have gone on and laid fourteen if necessary. We are quite happy, therefore, about laying the Terrenceville–Sydney Mines cable, provided that we are not delayed, and have to do it at a time when hurricanes such as "Hazel" are coming up the east coast of America.

I do not think that kinks are put in when cables are laid; kinks are put in when recovering cables or when making a splice and the ship is rising and falling in a swell. We have recovered many hundreds of miles, as have other cable ships, from depths of 3 000 fathoms. We can carry on picking up many miles of cable without kinks if the ship is not allowed to overrun the cable. When picking up it is possible to put a kink in at any time, if care is not exercised in the handling of the cable ship.

Mr. F. C. Wright: In Section 4.3 the authors say "Cost, fortunately, is a less important factor than it is in most other applications." I am not disagreeing with their opinion, but I am most definitely disagreeing with the words in which they express it. I would say that in a project of this magnitude and importance, and where the results of failure can be so dire, no engineer could possibly contemplate its engineering if he kept an eye on the cost of the equipment. The intrinsic value of many of the parts which go into these repeaters may be no more than a few pence, but it must be recognized that, in order to get a part worth intrinsically a few pence, it may be necessary to expend an equivalent number of thousands of pounds.

For that reason, the 5-step philosophy of the Bell Laboratories, which is so simply and ably expressed in the same Section of the paper, is more than fully justified. If there is one step in the five which perhaps in the final analysis is the most important, I think it must be the last, namely to develop the worker's faculty of self-criticism; because, however careful one may be in choosing the materials and processes, etc., in the final analysis, if they are not put together with the highest standard of self-criticism by the worker, it is inevitable that there will be some hidden fault or failure of the human element which will largely nullify many of the other steps which have been taken.

In Section 5 it is stated that the suitability of the torpedo type of repeater for use in very deep water has still to be established. On that point, I should like to say that, more than a year ago, two such repeaters were laid and recovered several times in over 2 000 fathoms of water. In one of these was included instrumentation of the type described in the paper, and in all respects the results which were obtained and the experience gained have given every ground for confidence that this design is appropriate for projects in deep water. During the past summer a further repeater has been laid and is to-day operating in 800 fathoms satisfactorily. I think that it can therefore be said that, while further experience is no doubt called for, all that is known to us to-day—and the successful conclusion of the North Sea cable to which Sir Gordon has referred is more evidence in the same direction—encourages us to believe that we should be ready and able to undertake a project of this type.

Mr. R. M. Chamney: I should like to say something which is of interest from an historical point of view; it has never come to light, because of the 1914–18 War. An old colleague of mine, now dead, Mr. C. Robinson, and I were working in the Post Office Research Department over 40 years ago and were dealing with the design and development of amplifiers. We had got quite a long way in 1913, and by early in 1914 the amplifier had reached a stage where we could quite confidently expect it to give excellent results under various difficult conditions. We actually designed a route for a cable from Scotland to Canada which would give three circuits, going from Scotland, probably via Shetland but possibly direct, to the Faroes, from the Faroes to Iceland, from Iceland to Greenland, from Greenland to Labrador and from Labrador to Newfoundland. Unfortunately, with the outbreak of war in 1914, the scheme never reached maturity; nor did we discuss it with the Americans; but I should like to put it on record that even then, 40 years ago, we were thinking of transatlantic cables of this type, and I believe that the project could have been quite easily carried out at that date.

With a coaxial cable and an infinite number of repeaters there can be an infinite number of circuits for an infinitely small cost; in other words, there are simply repeaters and no cable. Unfortunately, a submarine cable must be laid at the bottom of the sea, and must be recovered and relaid if there is a fault in a repeater. Some years ago I calculated the most economic distance for submarine submerged repeaters for a transatlantic or other long deep-sea cable, and I came to the conclusion that something of the order of 40 nautical miles was about the normal distance for their spacing. I am very pleased to see that the authors more or less agree with that.

Mr. D. T. Hollingsworth: There seems to be a difference of opinion as to whether kinking is caused during the laying or the recovery of the cable, and even whether it occurs at all if the cable is laid in a proper manner. The authors imply that it can be avoided if the tapes and other constituents of the cable

are all laid in the same direction. I should like to have their explanation of the mechanics of kinking and of how, if tapes are all in one direction, it will be avoided.

I should also like their opinion on the degree of hazard which exists if kinking occurs. I gather from the paper that, so long as the cable is laid in the correct manner, this hazard is not great; but the authors hint that it may be serious when repairing the cable, and I should like to know how serious they consider this hazard to be.

Mr. J. A. Smale: Since quite a lot of the discussion is centred on the question of the laying and possible kinking of the cable I would draw attention to the existence of a paper* which deals with this whole problem very fully. I suggest that some reference to this paper might be made, particularly as it is one of the few dealing with the subject of laying and recovering submarine cable, and it answers many of the questions which have been asked during this discussion. Kinks are not introduced during a steady pay-out, but only after the cable has been hanging at the bows or after a pick-up—conditions which might result during the paying out of the actual repeaters and the making of the final splice.

In Section 2.2 the statement is made that after 1908 the next important event in transatlantic communication was the establishment of a commercial radiotelephone service in January, 1927. Although the Section deals with radio circuits, it might be more correct to use the words transatlantic radiocommunication; even then, the statements are incorrect, because in 1926 the Marconi–Franklin beam circuit was opened between the United Kingdom and Canada. This beam circuit was put into commercial use for telegraphy, but it did, in fact, demonstrate the availability of a highly satisfactory radiotelephone circuit.

Mr. W. J. Archibald (*communicated*): In Section 8.3 of the paper the authors state that the feedback in the American repeaters is effectively removed over a very narrow band of frequencies. Will they indicate the magnitude of the dip in the $\mu\beta$ characteristic and the actual minimum value of $\mu\beta$ reached at the frequency at which the dip is centred? One would expect a large dip in $\mu\beta$ to affect the phase margin of the amplifier, while at the same time the actual value of $|1 - \mu\beta|$ at the centre frequency should approach unity if variations in μ, due to ageing of valves, are to be measured accurately at the end of the system.

Mr. J. J. Gilbert (*United States: communicated*): The authors have clearly indicated the difficulties and risks involved in the repeatered transatlantic telephone cable. Success will be attained only by adhering closely to the accumulated engineering experience and by use of the utmost care in the manufacture of repeaters and cable. Cable manufacture will be of unprecedented difficulty. Perhaps the most interesting requirement on cable performance is a consequence of the need for impressing on the cable insulation potentials of the order of 2 000 volts, in order to supply power to the repeaters. Until comparatively recently it was considered unsafe to employ much more than 100 volts on long submarine cables. This restriction was based in part on sad experience in applying a high voltage to one of the early transatlantic cables; but there was also realization that joints in insulation were not as strong electrically as the insulation itself. In the polyethylene cable, jointing has been studied intensively with the aim of obtaining a molecular bond rather than mere physical contact between the surfaces to be joined. The joint is then as strong and as lasting as the insulation itself. To gain still further assurance, the number of joints is reduced to a minimum by insulating the conductors in unusually long lengths.

Assurance of permanence and continuity of insulation and conductors presents many problems to the cable manufacture and there are others relating to the need, indicated in the pap of achieving throughout manufacture predetermined electri characteristics to an unusually high degree of precision a uniformity.

Mr. H. V. Higgitt (*communicated*): I note with interest that t diameter of the sheathing wire of the deep-sea cable is 0·086 This would be classed by the submarine-cable engineer No. 14 wire (B.W.G. No. 14 is actually 0·083in), as distin from the No. 13 wire now usually specified to be 0·099in diameter (B.W.G. 0·095). This difference in diameter betwe No. 13 and No. 14 wire might at first sight seem unimporta but when account is taken of the fact that torsional strength proportional to the cube of the diameter, it can be shown be very important. The importance of this is dealt with in t paper* by Besly and myself quoted by Mr. Smale.

This choice of such a small sheathing wire for such a hea cable (with well over half a ton per naut of copper alone), surprising. The weak torsional strength of such a cable is no important significance when paying out cable, provided th the process is continuous, but is very objectionable duri stoppages in that process and during any subsequent rep operations. It would be interesting to know what guided t designers in the choice of sheathing size for the deep-sea cab for it would seem that either they were not fully aware of t existing state of knowledge of the art or that other consideratio made it necessary to accept the disability of low torsional streng

Mr. K. Konstantinowsky (*communicated*): I have no experien with deep-sea submarine cables but quite a lot with steel rop for deep mine shafts, and I believe that there is a similar between the behaviour of the two. The submarine cable contai a central part—the cable itself—which does not contribu much to the mechanical behaviour of the finished cable and t armour which takes practically all the load. The same appl to steel ropes. I have in mind the inner part of such rop consisting of a hemp core which takes no load and the st wire strands laid up around it. It is common experience th steel ropes must always be under tension, otherwise there is t danger that kinks will be formed.

There are, to my mind, two factors contributing to the fo mation of kinks in steel ropes. One is that a twist is put into t steel-wire strands during laying up, and the other one—mention in the paper—is that variations in tension on the rope tend twist and untwist the rope structure. The rope can be handl much more safely if the twist in the steel strands is reduced to minimum. This can be done during manufacture if the stran are bent before final laying up in such a way that they assume t form, or nearly the form, of the spiral they will have in t final rope. There is not much that can be done, except t exercise of supervision and care, to make sure that the tensi on the rope does not vary too much.

Much effort is spent on controlling the tension on the su marine cable while it is on the cable ship, whereas it is impossi to do so with that part of the cable in the water, althou variations of tension may occur, owing to the formation of t sea bed or during a rough sea. It appears that nothing has be done to reduce the twisting of the submarine cable by applyi methods in the armouring process similar to those used duri the laying-up process for steel ropes. Is this intentional? C the authors give any explanation?

The authors state that buckling may occur if for any reas the twisting is localized. It seems reasonable to assume that would be harmless if reduced to a minimum, because dangero localization would be prevented. Will the authors express opinion on this?

* BESLY, J. C., and HIGGITT, H. V.: "The Recovery of Deep-Sea Cable," *Journal I.E.E.*, 1933, 72, p. 160.

* *Op. cit.*

Mr. A. L. Meyers (*communicated*): The formation of kinks during deep-sea operations is particularly liable to occur when picking up cable connected to a repeater. It is caused by the axial twisting of a helical layer of armour wires when subjected to longitudinal tension, and means for overcoming the difficulty will be of interest.

The tension in the cable varies from a maximum at the surface to a minimum at the bottom, but the accompanying axial twist may vary differently.* Except in ideal weather conditions, the longitudinal tension is not steady; thus the rise and fall of the cable ship in an ocean swell causes variations of tension, and waves of diminished and augmented tension travel down the cable, accompanied by corresponding waves of axial twist. The repeater, owing to friction against the sea-bed or to its inertia when suspended, interrupts these waves, with the result that under the relaxed tension near the bottom the cable may be thrown into turns which accumulate at or near the repeater. The augmented tension when the ship rises, or when the cable continues to be picked up, draws the turns into kinks. Breakage of the cable under the full tension as the kinks approach the surface may make it impossible to recover the repeater.

Twisting of a cable under tension can be substantially eliminated by the provision of two layers of armour wires, applied in opposite directions, with the lays so adjusted that an approximate torsional balance is obtained. Such double-armoured cable, however, is considerably more expensive than normal single-armoured cable. It might be thought that an appreciable length of double-armoured cable could be used in the vicinity of the repeaters, spliced to single-armour cable which would comprise the majority of the section between repeaters. The objection to this is that the sudden change in the torsional properties of the cable at the junction between the two types again partially interrupts the waves of axial twist, and there is then a tendency for turns to be thrown and kinks formed at or near the junction.

The suggested means* of overcoming the trouble is to use double-armoured cable next to the repeaters and single-armoured cable for the majority of the repeater sections, but to graduate progressively the torsional properties of the cable between the extremes of the two types, so that there is no large sudden change in torsional properties at any one place. Then the waves of axial twist will be progressively reduced in amplitude without serious partial interruption, and the tendency to throw turns and form kinks will be eliminated. The graduation of torsional properties can easily be achieved by continuing the wires of the single-armoured cable without alteration up to the repeater, and by applying over them a second layer of wires in the opposite direction, the number of wires in the second layer and/or their angle of lay with respect to the axis being progressively increased until substantial torsional balance is obtained in the double-armoured cable next to the repeater.

In a works trial the length that could be tested was naturally limited; the magnitude of each step in the graduation was greater and the distance between steps much smaller than would be desirable in practice. In spite of this, when tension was applied through a swivel at the single-armoured end, the double-armoured end being anchored and locked, a severe untwist per fathom in the single armour diminished smoothly through the tapered portion. Treatment which was effective in throwing turns and forming kinks in single-armoured cable did not do so with the tapered double-armoured sample. It is intended to carry out a sea trial of laying and recovery next spring, which if successful should also lay the bogey of the practical difficulties of using rigid repeaters in deep water.

THE AUTHORS' REPLY TO THE ABOVE DISCUSSION

Dr. M. J. Kelly, Sir Gordon Radley, Messrs. G. W. Gilman and R. J. Halsey (*in reply*): We have been honoured by having the discussion opened by Dr. Buckley, to whom the submarine-cable art has owed so much for many years. To Dr. Buckley and to Sir Stanley Angwin, who has repeated the contribution with which he opened the discussion in Chicago, we express our thanks for their comments on the background against which the project was launched and for their confidence in its success. We are sure that The Institution will take note of Sir Archibald Gill's suggestion that when the cable is complete there should be a joint meeting of The Institute and The Institution conducted over the new service.

It has already been made public that the overall cost of the joint project from the United States–Canada border to Oban is likely to be about £12 500 000; of this more than half will be spent on cable and repeaters between Oban and Clarenville. Without entering into the details of Sir Archibald Gill's calculation, we can say that all parties concerned are satisfied that the revenues should cover the operating costs and capital expenditure amortized over a shorter period than 35 years. In reply to Sir Leslie Nicholls, it is pointed out in the paper that separate cables for the two directions of transmission may be economic on long deep-sea routes with growing traffic requirements. It must be remembered that the power voltage required sets a limit to the number of repeaters that can be inserted in a cable, and therefore to the maximum frequency transmitted.

Everyone interested in transatlantic cables has looked at the northern route—Shetlands, Faroes, Iceland, Greenland, Labrador (or Newfoundland)—and a number of detailed plans have been investigated, the first following the failure of the 1858 cable. The maximum distance between landings would be less than 1 000 n.m., and the route as far as Iceland would be easy. The two western sections, however, would be difficult to lay and maintain. In view of the ice conditions, cable could be laid between Greenland and Newfoundland only in the late autumn, when the days are short and the weather at its worst, and all-the-year-round maintenance would be possible only if an ice-breaking cable ship were used. There is also considerable hazard from grounding icebergs.

Sir Archibald Gill's comments on cable ships are particularly welcome at a time when thought is being given to the changes required in their design and in that of cable machinery in order to meet future requirements. Captain Leech's long sea-going experience has rightly caused him to draw attention to the extreme conditions under which ship and machinery will have to work. Mr. Hartley has made considerable progress in the task of getting large rigid repeaters through the cable-laying gear. Solutions to the mechanical problems are needed quickly, because the usefulness of this type of repeater is in the immediate future.

The problem of component reliability was underlined by Brigadier Harris; an added factor to be taken into account is the way in which the quality of commercial products manufactured at different times tends to vary. While we, as engineers, agree with his remarks on "over-ambitious" systems, we must add that the desire to give maximum circuit provision must inevitably influence the design of future systems as it has that of the present main Atlantic crossing.

Mr. Dean has rightly stressed the importance of the chemist and the physicist in an undertaking of this kind. The processes

* Besly, J. C., and Higgitt, H. V.: *Journal I.E.E.*, 1933, 72, p. 160.

* British Patent Application No. 29795, 1954.

leading to ultimate failure may well be of a chemical nature, e.g. electrolysis, increasing in magnitude as the chemical products of such electrolysis increase.

We agree with Dr. Smith that a solid conductor would be desirable for transmission reasons; modern brazing techniques and methods of manufacture and inspection of the centre conductor increase the likelihood of obtaining a mechanically dependable solid conductor. In the past the use of soft-soldered joints and the possibility of hidden inclusions in the copper have led to the choice of a composite structure, and the present project has been based on designs which have proved their reliability by long usage. The possibility of water entering the repeater along the centre conductor has been guarded against in the British repeaters by special castellated seals and in the American repeaters by bonding the plastic seal to the centre conductor as well as to a coaxial tube. Vistac is polyisobutylene with a molecular weight of about 5 000.

A number of speakers have referred to the problem of cable kinking, and here we cannot do better than refer them to the admirable paper mentioned by Mr. Smale and others, and which is too well known to have been overlooked. The relative lays of the various parts of the cable structure do not materially affect its liability to kink but are important in restricting damage. We can assure Mr. Higgitt that the cable designers have for many years been concerned with torsional rigidity, both from the standpoint of recovery and of preserving the integrity of th coaxial structure. The theoretical design has been verified b extensive tests, both in the laboratory and at sea. Prototyp cable has been recovered from 2 500 fathoms with no imper fections and, on recovery, it coiled easily and cleanly.

Mr. Konstantinowsky refers to the similarity between cable and steel ropes; the lay of cable armour, however, is very muc more permanent than that of the strands in a wire rope and th tendency to unlay without tension is small. Preforming has, i fact, been tried on several hundreds of miles of deep-sea cabl without any noticeable improvement in its tendency to kink Kinking occurs whenever cable has torsional stresses withou adequate tension to overcome them; the worst conditions occu on starting to pick up cable in deep water.

Dr. Smith has referred to the mounting space in the two type of repeater. In the British repeater with dual amplifiers, abou one-third of the available space is occupied by directional equip ment. In reply to Mr. Archibald's question, the shunt due to th crystal at resonance is such as to reduce the feedback by nearl 30 db, to about 2 db. The feedback-network design is such tha the reduction in phase is held to a maximum of about 2° just belo resonance and the phase at and above resonance is increased.

In conclusion we are grateful for the contributions made t the discussion by our American colleagues, Mr. Duncan an Mr. Gilbert.

DISCUSSION ON
"POST-GRADUATE ACTIVITIES IN ELECTRICAL ENGINEERING"*

Before the NORTH MIDLAND CENTRE *at* LEEDS, *9th December*, 1952, *the* SOUTH MIDLAND CENTRE *at* BIRMINGHAM, *5th January*, *the* SOUTHER CENTRE *at* SOUTHAMPTON, *21st January*, *and the* NORTH-EASTERN CENTRE, *at* NEWCASTLE UPON TYNE, *23rd March*, 1953.

Prof. G. W. Carter (*at Leeds*): The authors are quite clear that they regard research in universities by newly graduated men as not so good a training, nor so effective a mode of obtaining scientific results, as the same research done in industry. I have had the opportunity to discuss this problem with the directors of certain research institutions, and, on the whole, I did not find that they took quite so sweeping a view as the authors; the general opinion was that, for a few really research-minded men, research work in universities is beneficial. The practice of my department is based on this view, so that we accept a few research students, but take care before doing so that they are really suitable.

It is worth pointing out that when a graduate elects to remain at the university for a further year or two, he does so at some financial sacrifice, since the level of research grants is over £100 a year lower than the level of remuneration for graduate apprentices. The research students make a valuable contribution to the life of a university department, forming as they do a link between the teaching staff and the undergraduates. When members of the teaching staff are themselves engaged in research, it is very useful for them to have one or two research students working on allied problems; in suggesting that senior members should continue personal research, the authors are perhaps forgetting that it is difficult for them to do so in isolation and without the complementary research of younger men. It is true that the research of a young graduate benefits the man more than the industry, but that does not discredit it.

I believe that a strong case can be made out for saying that young graduate may benefit greatly from a period of research i the university. It gives him something that he has not ha during his undergraduate course—time to think and consolidat his knowledge. In industry people fulfil specialist function and the man who carries out tests is frequently not the same a the man who plans them; in the university the young man ha the better training of doing both the thinking and the exper menting himself. Furthermore, the specialists in industry ar extremely expert in some fairly narrow field, whereas universit teachers are not usually experts, but are able to tackle a wid variety of problems from first principles; I believe that it may b more stimulating for a young man to discuss his work with senior who is not too expert, especially as in some problem progress depends upon finding that the ideas of the expert wer partly wrong. For these reasons, I consider that research wor in universities may be a training of great value. Of course, would be better if we could get our research students when the have had a period of industrial experience. If, besides puttin into effect the research scheme described in the paper, th authors can also use their influence to enable a few men to spen a further period in the university at the end of their apprenticeshi such men might benefit greatly.

Dr. E. C. Walton (*at Leeds*): The authors are opposed to th creation of new technological universities for the purpose enumerated in the first paragraph of the paper. With th opinion I agree, but I am unable to accept the implication in th last few lines of the paragraph to which I have referred: I cann believe that the authors would deplore any movement by existin

* GIBBS, W. J., EDMUNDSON, D., DIMMICK, R. G. A., and LUCAS, G. S. C.: Paper No. 1265, February, 1952 (see 99, Part I, p. 161).

universities and colleges to provide "many more graduate engineers" and to afford "extensive facilities for research and development in all branches of industry."

In my view there is a case for the improved training of selected men from the group pursuing the normal industry-based route, i.e. a five-year apprenticeship coupled with part-time courses leading to Ordinary and Higher National Certificates. This improved training should take the form of industry-based sandwich courses of the type referred to by a previous speaker. The general plan for such a course might be: one year in industry with concurrent part-time attendance at a technical college followed by selection on merit for a subsequent four-year sandwich course (6 months in college, 6 months in industry per annum) and finally a further year in industry coupled with post-graduate training. This type of course offers a wider and deeper training than is possible with existing National Certificate schemes. The existence of such courses should encourage boys from secondary grammar schools who might otherwise be attracted into other industries, to enter the electrical engineering industry. Such sandwich course schemes are dependent upon a large measure of co-operation from industry, as the students will need a maintenance grant or the payment of wages during each period of six months spent at a technical college.

The need for such sandwich courses giving improved training facilities to selected men from the industry-based stream has, I think, been emphasized by the authors, particularly by the conclusions given in the last paragraph of Section 5.

Turning to post-graduate evening or part-time day courses at technical colleges, the weakness mentioned under (*a*) in Section 4.2.2 may be overcome by arranging for a member of the full-time college staff to act as organizing lecturer, responsible for co-ordinating the work of the several visiting lecturers.

The interwoven series of lectures given in the Advanced Engineering Course described by the authors appear to be a most valuable form of post-graduate training, with immense possibilities. Can the authors describe the method of selection employed for entrance to the course? Is the course intended mainly or solely for those graduates likely to be engaged on research or development work?

The content of the Advanced Design Course (Section 7.3) appears at first sight to be unsuited to the needs of men with Higher National Certificates. The severely practical bias given to the lectures appears to be unnecessary for men who have presumably followed a five-year apprenticeship in the works. It may be, however, that the lectures are concerned with new developments in manufacturing techniques, rather than with standard methods. Perhaps the authors would enlarge on this point.

Mr. A. J. Coveney (*at Leeds*): There is no doubt that the burden of scientific training has been for years the responsibility of the industry, and the paper indicates the advanced nature and the costliness of such training which has been undertaken by a large manufacturing organization in the electrical industry. Will the authors express their views on the merits or otherwise of setting up a specialized National College or Institution, sponsored by the Government and supported by both the industry and the universities, to give this advanced training to the electrical engineer? Such a college has already been set up by the aeronautical industry at Cranfield, which, in addition to providing the post-graduate course, also provides the facilities for advanced technical research and training.

There has been no mention of the Faraday House student. Possibly, in view of the sandwich course, the authors would recommend a different syllabus in the scheme for these students.

With regard to interchange, I think not only should manufacturers send their senior technicians to give lectures to colleges, but also, as mentioned in Section 4.2.3 of the paper, by the lectures already given by Prof. Carter, the colleges should likewise provide certain lectures to students at manufacturers' works. I would also stress the importance of exchange of students between the manufacturing industry and the supply world and vice versa. This interchange of knowledge with the various problems and difficulties concerning each section of the work would help to solve many of the problems of to-day and to reduce some of the delays in both planning and production.

The authors refer to eight professorships and ten doctorates awarded in the last ten years to men from the industry. It would be interesting to know what sort of training these men received.

The immediate problem is no doubt the acute shortage of suitable young men. With the demand for technicians in the Armed Forces and attractive appointments in the Colonies, British industry must provide suitable and equal prospects both in advancement and money, if it intends to retain its share of these younger men, and no doubt with the rising cost of living and taxation, the financial aspect is considerable.

I would conclude by quoting the well-known old Chinese proverb:

> If you plan for one year, plant grain.
> If you plan for ten years, plant trees.
> If you plan for a hundred years, plant MEN.

Prof. A. Tustin (*at Birmingham*): There is urgent need, with no fear of surplus, for a greatly increased flow of men with the knowledge, the keenness and the personal qualities to press engineering development ahead faster. The honours degree standard, supplemented by two years of practical training, is only a beginning of the process of building up the knowledge and skill that these men require. They require, in general, more mathematics, and in particular a deep and intensive study of different sets of specialized topics close to their particular specialization. It is agreed that such men should be thrown at a very early stage into work that calls for creative and discriminative mental effort, and for the play of their highest qualities of intellect and determination.

The controversial issue that is raised is whether such men would progress better in the long run, as a first alternative, by returning to the university for one, two or three years as post-graduate students, or as a second alternative, by being employed on development work in industry for the same period. It is important to note that the period in question is most usually only one year, and only for exceptional students should two or three years be contemplated.

I suggest that both of these courses are inherently capable of giving the required result provided that certain conditions are met, and that we must change the question entirely from the form "whether A or B" to the form "how can we make both A *and* B effective?"

I am well aware that post-graduate training in universities needs improvement; but perhaps even more so does post-graduate training current in industry. Let us find out what is wanted in both fields. In the first place, as Prof. Willis Jackson said in the discussion in London, there must be no question of doing away with research in the electrical engineering departments of the universities. Industry's first demand on the universities is for first-class undergraduate tuition. It is, in my view, essential in the long run, if we are not to see the standards of undergraduate teaching sink to squalid incompetence, that a high level of research activity should be maintained in such departments.

If research work is being properly carried out in engineering departments of the universities (and industry could and should help to see that this is the case), there is scope for training

there a limited number of post-graduate students. At present, because, as the paper admits, industry is unwilling to release men to return to the universities, a very large proportion of such places as are available is being filled by foreign students, in some instances much more than 50%. Thus industry is throwing away a resource that is at its disposal.

With regard to the training that can be given to a young man in industry, the authors should make more clear whether they mean their remarks to be taken to apply only to the organizations with which they are associated, where they may be justified, or to industry as a whole, where I am sure they are not. The authors know quite well that in very many departments in many firms the real need is to get young but technically mature men into the departments who have a fresh outlook and have drunk waters of inspiration from some other fountain than the somewhat stagnant pool of the department's own inbred resources.

For most of an engineer's life, of course, his further studies must unavoidably be in parallel with his tasks of production, with the daily pressure of delivery dates, and the necessities of doing the job by the tools already to hand. The university gives a young man something else—it gives him time. Time to probe his problem to the bottom, to try unconventional approaches, to read widely rather than narrowly, to draw ideas from discussion with a great variety of persons.

To sum up, I would say that I believe the author's scheme of training within industry is of excellent intention and of great potentialities, but even in their own exceptional organization it will need immense devotion and effort to supervise it so that it will give anything like the results the authors hope for. I beg them also not to forget that, somewhat unsatisfactory as they think post-graduate training and research in the universities to be, it should be appreciated nevertheless that it is of overwhelming importance to the electrical industry that the work of university departments should be supported and strengthened, and that this requires the continued support of extra-mural research and post-graduate work. It is not at all a question of the research requirements of industry being turned over to the universities; such an idea would be ludicrous. The resources of personnel and materials of the university electrical engineering departments are trivial compared with those of the great industrial and Government research establishments. The point is quite a different one, namely that their proper—if necessarily quantitatively small proportion of the total—research effort must be maintained and fostered within the walls of the universities.

Mr. L. L. Tolley (*at Birmingham*): There is a problem which I think Prof. Tustin describes under alternatives "A" and "B," and this is a personnel problem which is not capable of a definite solution. One can quite see, as Dr. Gibbs has pointed out, that if the individual remains at the university and does not get into industry until he is about 30 years old, he gets in a way set into a groove; he has not been called upon to take responsibility for getting results, and if he does not do that until he is well on in years he may never get to exercise that responsibility. On the other hand, it is perfectly true, as Prof. Tustin has said, that to cut short the best men's post-graduate studies arbitrarily might be in certain cases a serious mistake. The problem seems to be not a question of whether it is alternative "A" or "B" in general, but for each individual, and I see no option but to suggest that the university must give the indication whether the individual is such a person as ought to go on for several years.

Prof. Tustin also referred to a point which occurred to me, that perhaps the organization with which the authors are associated has rather special conditions; it is a very large concern and it must have a large say in the technological training arrangements in Rugby. That does not apply to the same degree to firms in larger towns, nor to small firms in small towns. The Post Office is almost at the other extreme; it has a very small number of posts filled by men who might be taking post-graduate courses such as those mentioned in the paper. In Birmingham we should have perhaps six or seven, in Nottingham, Leicester and Coventry there would be only one or two, and none at all in the smaller towns. Obviously we have so few people that we cannot have the technical training establishments in the towns where the people are, and we are bound to depend upon the universities, apart from the fact that we do operate a National Training Centre for specialized courses on types of work with which the Post Office is particularly concerned. I am sorry I cannot quote comparable figures, but so far as I can remember the growth of technical staff in the Post Office has been very similar to that indicated in Fig. 1. I think the percentage quoted by Dr. Gibbs would apply fairly well to us also. I was very interested by Fig. 2. We have a somewhat similar system of grading, though not precisely comparable, since ours is a coarser grading dividing the men into five groups instead of nine. The main difference which would prevent a precise comparison is that we are really grading against "good average," which means that as the new man who comes in is not then "good average," our new man goes to the "C" end, whereas the new man in Dr Gibbs's grading goes into "B." I wonder whether that big "pile up" in "B" consists to any major extent of the new entrants.

Mr. J. H. Patterson (*at Birmingham*): Mention was made of the "finished product." We have a training scheme, and I have consequently had fairly considerable experience as regards this "finished product"; I have been very disappointed indeed. Whenever a university graduate has been sent to me for a period of training, to learn perhaps more of the practical side than he has done hitherto, I always invite him to my office in order to outline the general run of what is intended and what we are trying to do. He is questioned to find out where his inclinations lie, and every endeavour is made to give him as much experience as possible in that direction. It seems to me that these people are "spoon fed" too much. My experience has been that they lack initiative to a very marked degree. The only man, in a matter of dozens, who showed any initiative was a Greek. He came and asked questions and I was very pleased indeed to do everything possible to put him on the right track. At the end of the period of training he was most grateful. I have found generally both a lack of initiative and inability to search round and find out information for themselves.

I think a great deal depends upon the ability of the teacher himself. The teacher and mode of teaching is of the utmost importance. A poor lecturer can spoil a good student, and conversely, a good one can produce results from an indifferent type of pupil. I suggest that more initiative should be inculcated into the students so that it will be forthcoming as and when they go out into business.

Mr. W. H. Bell (*at Birmingham*): On this question of research, the doubt is how far people will find suitable leaders under whom to serve their research apprenticeship in industry. (It has been pointed out that the authors' organization is exceptional in this respect.) I spent a good while in industry myself and on returning to the university realized with surprise how far the standard of research in industry is liable to fall below the best quality.

It is the function of the universities to produce ideas, and not prototypes, though if some places in industry also produce ideas that is all to the good. The Percy Committee suggested that whereas our ideas are good the difficulty is in getting them put into practice. I do not see that it would help to transfer the potential source of ideas from the universities to industry, even if its fruitfulness were not diminished in the process. There might

On the other hand, be a case for a really joint effort between universities and industry by analogy with the position of the teaching hospitals in medical training. This joint effort would presumably take the form of a research institute, and the possible function of such an institute has first to be established economically as well as academically.

Given the present facilities, I think it is a fair compromise that the most promising students should do a one-year research course (e.g. the M.Sc. course which is available at a number of universities) to start their apprenticeship in research. This gives an opportunity for vocational specialization of studies and also gives the student an opportunity to exercise the initiative which the authors find so frequently lacking in first-degree students. I think such a one-year research course is the biggest contribution which the universities can make to the needs of industry at the present time, and I hope some of the graduates from such courses will eventually apply their abilities to the problems of production as well as those of developments.

Dr. E. H. Norgrove (*at Birmingham*): I like Prof. Tustin's treatment of "squalid incompetence" in teaching. I agree that research is no substitute for practical training, but under his eminent predecessor Prof. Cramp, it was possible to work on research and at the same time incorporate a good deal of practical experience.

I should like to ask who are these supermen who do the assessing in the organization we are discussing, how are they trained and who assesses them? Finally, all this talk about education and so forth has been on a very high idealistic level, but we have to remember that the man has to live his life as an individual. It is all very well to say that he must spend so many years doing this and so many years doing that, but a man wants to get married, have a home and children, and I have yet to see any reference to that in these rather grandiose educational schemes.

Mr. A. R. H. Thorne (*at Birmingham*): Little reference has been made to the financial side of training. It appears that the assessors of the usefulness of various students to industry are interested only when the fruit is ripening; what about the green state, when the students are given early training?

Those who are fortunate enough to obtain substantial financial aid are able to go to a university. There must be thousands of students who only attain Higher National Certificate standard owing to restricted financial aid, whereas help at this vital stage would enable a large proportion of them to obtain a higher standard. I am interested to know whether active steps are being taken within industry to help those students who would benefit by additional training, this help being apart from scholarship aids of the usual kind.

Mr. E. S. Hall (*at Birmingham*): Some of the views put forward have been rather weighted to suit the ideas of the authors; for instance the curve in Fig. 1 could well have been drawn as a straight line. Similarly I feel that the curve of grading of the pass-degree engineer as shown in Fig. 3 looks rather a different shape if you miss out the man who achieved Grade A+ at the age of 30. It particularly alters the line drawn at the front part of the curve. You must also remember that these points on the curves represent different engineers and not the progress of one engineer.

I should like any information the authors can give as to how these curves would look for the commercial engineer and for the engineer who takes up the manufacturing side. It has already been pointed out that we in this country do not suffer from the lack of ideas but do suffer seriously from putting these ideas into practical shape; that is where the manufacturing engineer comes into his own, and it is therefore very important to attract good men to the manufacturing side. I should like to know whether the curves for the commercial engineer and the factory engineer would differ markedly from those for the design engineer; in other words, are the same qualities and the same type of individuals required for the design engineer as for the factory engineer, and does the honours man, in general, still show up markedly superior for these applications as compared with the pass-degree and the Higher National Certificate man?

Mr. W. S. Terzi (*at Birmingham*): The authors state that when a man has completed his university course they do not consider it desirable that he should stay on for research, and when he has once gone into industry his employers are not prepared to release him or, they say, he is not prepared to leave to go back to the university, mainly for financial reasons or considerations of seniority. The question, then, is when is he going to do research if that is his wish? They have examined the problem from their angle, but they have not really looked at it from the point of view of the person concerned. We are living longer these days, why cannot we be left to enjoy our youth? Why all this concern over the age at which we achieve executive status?

Industrial and university research are usually not on the same level, nor have they had the same objectives; the former is usually a means to an end whilst the latter is an end to itself. If industry will not release men to go back to the university, it may become impossible for the staff therein to carry on research work. It is not unlikely that this may result in the university staff becoming static and unprogressive, and that with tragic consequences as Prof. Tustin has already pointed out. Such a state of affairs would also obviate the paramount factor which differentiates between the individual universities from the technical point of view, this factor being the head of the department and the various research activities in which the staff are engaged.

The suggestion made by Mr. Bell that some students be allowed to stay on for one year after graduation is excellent. As an undergraduate the student has to cover a certain curriculum; if he is allowed to stay on for one year after his finals, then notwithstanding the technical self-confidence which he would gain from such a period, the fact that he will come into direct contact with the staff in the research department may have a great influence on his future. Dr. Gibbs has said that a student may work on present-day research at the university and then take an appointment which is not at all related to that research. That may happen, but on the other hand a man's future may be changed owing to his post-graduate activities in the university—his coming into contact with a certain type of work which he may decide to continue afterwards. I think that this is a point to be emphasized.

Mr. P. A. White (*at Southampton*): As a student I find the post-graduate training scheme rather discouraging. It seems to me that anyone who wants to get on in the electrical industry is faced with many years of extremely difficult and exacting work, in fact he is faced with the prospect of being a glorified schoolboy until the age of 35.

This prospect might discourage many students from entering the electrical industry. The scheme may well produce first-rate technical robots, but I wonder whether the authors have ever considered what type of men it will produce.

Mr. A. T. Crawford (*at Newcastle upon Tyne*): I agree with the authors on the desirability of practical training and on the undesirability of the creation of new institutions. It would appear that graduate engineers, on completion of their university training, are not particularly anxious to receive practical training on a course. They are more inclined to look solely at the immediate financial reward they can obtain for their services, and they may be offered appreciably more than would be paid to an engineer undergoing his two years' post-graduate training.

One speaker has suggested that the authors wished to eliminate research work at universities, but I do not believe this is their

intention, and I consider it essential to retain research facilities at universities, with an honours school having a limited number of post-graduate students. The limited capital expenditure that can be incurred necessarily limits the useful work by post-graduate students.

The authors mention a universal figure of expansion in the industry of 4–5% relating to the technical staff necessary to maintain conditions on a satisfactory basis, and subsequently talk of engineers trained in England taking up appointments abroad. Is this figure truly universal, or is it only an average with comparatively wide variations in different countries?

I endorse whole-heartedly the sentiment that grammar schools and public schools should be convinced of the soundness of recommending some of their best men to take up electrical engineering as a profession, and the universities can assist in this by allowing their staff to teach engineering subjects in grammar schools so that the entrant from a grammar school comes with a knowledge of more than simply pure science.

Various comments have been made on the level to which mathematics is taught in different institutions, and the authors have found widely different standards in their men who should normally be at the same level. There is no suggestion that the man on the lower level could not absorb the more advanced work, and more vigorous steps should be taken to improve the level of mathematics up to the highest standards now taught in engineering courses throughout the country; I feel also that much more attention could be devoted in the earlier years to advancing the level of mathematics taught in primary and grammar schools. There would then be no need for the men in later life to spend time on further training in the subject.

On the charts shown in Fig. 3 it is not quite clear why there should be such a limited number of honours graduates after the age group of 35, and the authors say that the evidence they present may lead to erroneous deductions unless it is analysed very carefully. For those of 35 years of age and over, pre-war technical education was reasonably stable, and I thought it would have been possible to gain useful evidence from the charts. It appears, for example, that honours graduates with few exceptions have maintained their early promise in later years, whereas pass graduates have more nearly remained as average engineers.

Dr. W. J. Gibbs and **Messrs. D. Edmundson, R. G. A. Dimmick** and **G. S. C. Lucas** (*in reply*): In the discussion, university professors and other educationists have raised the question of research training in universities. We have to differentiate clearly between those men who intend to make careers in industry and those who wish to spend much of their lives in research. In general, it is undesirable for a man who seeks a career in industry to remain at the university after graduation. We agree that there are a few exceptions to this general statement. It is said that unless university staffs have post-graduate students working under them, their own research work must suffer and the teaching of undergraduates must ultimately deteriorate. We think that if the universities retain those seeking careers mainly in research plus the few exceptions cited above, they will have enough post-graduate students to prevent any deterioration in either teaching or research.

It is our experience that those who do stay at the universi immediately after graduation for two years' research are n willing to follow this with an apprenticeship in industry. The expect to by-pass the practical training; they consider their tw years' research as an equivalent, which it emphatically is no With regard to the period immediately following apprenticeshi we have found that men are reluctant to return to the universit at this stage because they are keen to take up an appointmen either with the firm at which they were apprenticed or elsewher They are even more reluctant to return after some years on th staff because by then they are specialists, already immersed i their own researches, and they fail to see what benefit they wou obtain by breaking into this work to return to the university.

Let us repeat that we do not ask the universities to abando research work and that we recognize that exceptional men hav to be treated as exceptions. However, it does seem undesirab for a man aiming at an industrial career to be kept, after gradu tion, in what is almost admitted to be the leisurely atmosphe of the university where he "has time to think." Mr. Terzi sa "why cannot we be left to enjoy our youth?" We ask educ tionists to ponder this question and its context. We also dra attention to the contribution of Mr. White, another student, wh is dismayed at the prospect of additional study in industry. N one is compelled to undertake further academic study aft graduation, and many do not.

We can now deal with some of the specific questions raise Dr. Walton raised some concerning the Advanced Engineerin Course. It is intended only for those wishing to take up desig or research, and the selection is made from an assessment of th work done in the preliminary course at the College of Technolog and the examination that terminates it, and as a result of ind vidual interviews. The lectures in the Advanced Design Cour are concerned with special techniques and new developments.

In reply to Mr. Tolley, the engineers graded "B" form th largest group because all new entrants start with that gradin Mr. Hall suggests that Fig. 1 could have been drawn as a straig line. This is true. No set of points determines the nature of th curve to be drawn through them. We chose a compound-intere curve because the industry is expanding according to this law an it seems reasonable to suppose that the expansion of technic staff follows the same law. However, if a straight line is drawn, reinforces our argument that a vast increase in the number graduates in engineering is unnecessary.

Mr. Hall is right in suggesting that the curves of Figs. 2 and would be a little different for commercial and factory enginee For instance, personal qualities are of prime importance in t commercial engineer. While, therefore, the grading of honou graduates would not be markedly different, the pass men an Higher National Certificate men would tend to rise more quick A more important difference lies in numbers; the factory an commercial departments tend to have a smaller proportion graduates on their staffs.

We regret that owing to space limitations we cannot discuss the points raised, but we are very grateful to all who have tak part in the discussions and thereby given us all food for thoug on this most important subject.

Abstract No. 1848
Mar. 1955

EAST MIDLAND CENTRE: CHAIRMAN'S ADDRESS

By J. H. MITCHELL, Ph.D., B.Sc., F.Inst.P., Member.

"OBJECTIVES IN INDUSTRIAL RESEARCH"

(Address delivered at LOUGHBOROUGH, *12th October,* 1954.)

It is the custom for your Chairman to take as his theme some major problem or feature of that section of the electrical industry with which he is directly concerned. I should, by this precedent, talk about organization for research, but I find some difficulty in doing this. Any attempt at organization of research must in large measure depend both on the structure and policy of the firm or institution concerned, and on the personality of the man responsible. In my view the subject is best left to a discussion group.

In considering organization for research it is probably more nearly true to say that it is the research that organizes the planners. No major project in a research laboratory can be guaranteed to follow the path expected, and I intend to describe such a project which has not yet achieved its aim and yet in its pursuit have been found prospects of good commercial returns.

When I joined my present firm, seven years ago, the major project facing me was the design of an electronic telephone exchange, and a good glimpse of this possibility was given to you in Mr. Flowers's paper in 1952. Before the war an all-electronic exchange was considered a futurist dream, but seven years ago people were talking of it as being ten years away. However, to be a practical proposition the electronic exchange must compare favourably in first cost and in subsequent maintenance over its working life with the electro-mechanical.

In a mechanical exchange the speech path must traverse many metallic contacts. These are by no means ideal devices for carrying speech signals, they are affected by dust, atmospheric pollution, mechanical wear, and in inductive circuits by electrical wear, and unless an exchange is carefully maintained steady deterioration will occur. By 1947 it had become clear that if we were considering rapid switching, electronic devices would beat mechanical in total number of operations, but if the operations were spread over a long period the position would become more complex.

Before proceeding to considerations of greater detail I should point out that the existing British automatic telephone system is a step-by-step arrangement in which the same switch is used for determining the path of a telephone call through the exchange as well as making the speech connection. In some Continental and American systems control of the route taken by the call through the exchange is separated from the switches making the speech connection. There are arguments for and against these systems. It is claimed that in the British system wear is distributed over the whole of the exchange where the speech switches also have to search, whereas in the other system the wear is kept to the control part of the exchange where it is less likely to cause deterioration of the speech circuit. At the present moment the overall cost, which includes first cost and subsequent maintenance, and general performance are so closely equal as to make a choice of either system extremely difficult. I must mention, however, that a mechanical crossbar switch having low wear and high reliability, which shows to advantage in speech circuits, has recently been evolved and is used very effectively in Europe and America.

In considering the electronic exchange there are three possibilities: to add electronic circuits to existing exchanges wherever advantageous; to evolve an electronic system on the American and Continental lines in which electronic circuits replace the control equipment while the speech paths are still electro-mechanical; to design a wholly electronic system with a new approach such as that of the time-division multiplex described by Mr. Flowers.

Reviewing such possibilities seven years ago there appeared to be many lines of attack, but we invariably come to a major stumbling-block, namely long-life valves. A telephone exchange lasts at least 25 years, and valves with many tens of thousands of hours' life would be necessary. Moreover, the heater current of the ordinary valve, when used in the many tens of thousands which might be needed for a large exchange, could become a serious drain on local power supplies, and the resulting dissipation of heat would present a major problem in equipment and building design. Either a long-life valve with extremely low filament consumption had to be developed or some other solution found. There were two possibilities: cold-cathode tubes, which had been known for some fifty years, but were only then beginning to show promise of the stability required, or devices such as the transistor, then but a gleam in the Bell Laboratories' eye.

I came to the conclusion that it was necessary to carry out a detailed investigation into the possibilities of the cold-cathode valve. Simple tubes of this type have long been used as voltage stabilizers, since the voltage drop in a discharge path remains reasonably constant with quite large variations of current; many years before the war the trigger triode was developed, together with the hot-cathode equivalent, the thyratron. The cold-cathode trigger triode appeared to offer possibilities in exchange switching circuits, particularly if it could be made to carry speech, since it required only a small trigger impulse to close the circuit. Previous experience of my colleagues and myself indicated that in some tubes of a batch the noise level would be quite high, while in others it would be low enough for the purpose we had in mind.

Accordingly an investigation was started to obtain fundamental information on the practicability of a speech switch. We already knew that impurities and irregularities in cathode surfaces would cause the discharge to jump, with a noise burst at each jump; it was also known that various modes of oscillation could occur within a discharge, and that impurities within the cathode would cause rapid variation in electron emission, with consequent noise fluctuation. It had also been found necessary to keep the gas clean and pure to avoid cathode contamination and to secure a satisfactory life of the tube. We were not alone in this approach to the problem, which was also being considered on the Continent. Mr. J. R. Acton carried out a theoretical investigation which indicated that the noise might be reduced to 100 dB below the reference level of 1 mW and that the impedance of the discharge in the speech frequency range would vary inversely with the current. It is clearly desirable to work at the lowest impedance possible, to avoid crosstalk from other circuits in the exchange and other sources of interference.

Dr. Mitchell is with Ericsson Telephones, Ltd.

The first practical outcome of these early investigations was a diode; later a trigger tube was produced which had quite interesting characteristics. Unfortunately, in order to keep the attenuation through the valve to less than 1 dB, it was necessary to use input and output transformers, and this made the complete switch uneconomic. Consideration was given to the design of a tube which would provide a certain amount of amplification, but this was not then successful and we decided to suspend this investigation.

In the meantime we had been seeking the electronic counterpart of a uniselector switch, whose principal action is the selection of one of a number of outlets. A device of this type had been produced in a cathode-ray-tube switch. Mr. Acton sought to rotate a discharge round a circular anode, and thus the Dekatron was evolved. This tube, which is now becoming well known, is quite simple in action, but we were disappointed when we examined its use as a speech tube and again we put in abeyance this side of the investigation.

Thus far you will see we had failed in our objective, but an interesting position had arisen. We had first produced an extremely accurate reference tube in the diode we had investigated, and at the same time we had acquired a great deal of fundamental information on the operation of reference tubes and on stabilizer-type diodes. We had also developed a useful and relatively inexpensive trigger triode, and lastly we had evolved a new type of tube altogether in the Dekatron. Visitors to the laboratory became interested, samples were made, tried out by other laboratories, and later sold. The valve section started a miniature production unit and the whole grew and consumed valuable research space until finally we decided to build a new laboratory and production unit in which we could concentrate on the development and manufacture of cold-cathode tubes. The building which was constructed takes such care for its cleanliness, air conditioning, style and general working conditions that it has become the envy of the factory and other parts of the laboratory.

An offshoot of our primary objective had itself given birth to a new primary objective. It soon became apparent that further by-products were about to appear.

When I joined the company a service I implemented was to ensure that the factory and laboratory as a whole were adequately provided with test instruments, and we set up a small specialist instrument section to do this. During early meetings with some of our valve customers, the proposal was made that we should undertake the manufacture of certain specialized instruments, and in particular those required for nucleonic studies. This manufacture was started, and has grown so rapidly that recently we have had to move it into new premises of some 7 000 ft^2.

To return to the primary objective, an electronic telephone exchange: we have seen that cold-cathode tubes can be regarded as the children and nucleonic instruments as the grandchildren of the original project. Both of these offspring quickly grew to the factory production stage—but what of the electronic exchange itself?

Let us return to where we left the speech tube in abeyance. Three lines of development emerged from the work that had already been completed.

Although the speech tube in its simplest form did not appear to be an economic proposition, we had proved that the noise could be reduced to a practical level and that it could pass speech. A feature of the discharge—its negative resistance characteristic—had not been fully explored, and it might b possible to utilize this to allow resistive rather than inductiv coupling in the circuit arrangements. Work along these line continues.

In the London director system the first three digits dialle route the call through a complex network of trunk lines to th called exchange. This was always heavy work for a mechanica switching mechanism. The problem is much easier if the route i controlled by a cold-cathode director: equipment now on tes at an exchange is proving very promising. There are many othe similar applications now in an advanced stage of development.

In my opinion one of the most promising alternatives to a all-electronic exchange is the crossbar switch with electroni selection and control of the call through the exchange. In thi development there are a number of applications for which cold cathode tubes are ideally suited. For example, it is ofte required to operate the magnets by a very-low-power markin impulse, so that latest tube to be produced has a mean outpu power of about 10 watts, peak power of about 20 watts, an operating power of a fraction of a milliwatt, yet the tube size i quite small.

What of the wholly electronic exchange? It is still too costly and further component developments are still required. Perhap it may be that the transistor may ultimately prove reliable efficient and inexpensive enough to play an important part.

And so we have come full cycle. We are still some way fror solving our first problem, the all-electronic exchange. We ca see the way to design semi-electronic exchanges with cold cathode tubes playing a large part, and we can contribut materially to improving the existing type of exchange.

In arriving at this state we have founded a new industry, th manufacture of cold-cathode tubes; apart from our own demand for them, industrial, computing and business accounting market have welcomed them. Trigger tubes of greater stability, greate speed and higher power are needed and in fact are on the way and the Dekatron is increasing its versatility by greater speed o operation and is being redesigned as a shifting register tube As industry becomes aware of the high reliability and long lif of these new tubes, further interesting valves will be require and so our valve project must expand to supply industria demands, while still being mindful of exchange possibilities.

In the train of this comes another industry—the electroni instruments for nucleonic and industrial purposes.

Instead of talking about organization and planning for researcl I have taken you through one facet of what may prove to be th most interesting and most difficult phase the telephone industr has ever faced. The first phase was the simple exchange manuall controlled. The next phase was the introduction of automati exchanges, with the more recent introduction of trunk switching the third phase will see the use of semi-electronic or wholl electronic systems—no one can be certain which. In doing thi I have tried to reveal a facet that does not normally appear: it i that research has a knack of planning itself, and that the primar object can as it were give birth to secondary objectives, each o which can become a primary objective and so give rise to a second generation of primary objectives, and so on. In uni versity research the individual worker claims he must be free t go the way his inspiration takes him: in industry more tean effort is required but it can become a most uninteresting problen if it goes exactly the way it was originally planned. Commercia enterprise must follow the research in its development if it is t take advantage of the rich rewards research has to offer.

Abstract No. 1849
Mar. 1955

MERSEY AND NORTH WALES CENTRE: CHAIRMAN'S ADDRESS

By P. R. DUNN, B.Sc., Member.

"CABLES—SOME POST-WAR TRENDS"

(ABSTRACT *of Address delivered at* LIVERPOOL, 11*th October*, 1954.)

The uninitiated user, who generally takes cables very much for ranted, may sometimes wonder what the cablemaker has to how for his laboratories and for the money he spends on them.

The times demand, and have seen, active technical advances. The use of electricity grows exponentially; transmission voltages ontinually rise; telecommunication frequencies have increased thousandfold in the last decade; and new cable materials and echniques have been developed and exploited to meet these onditions. It may be of interest to survey what the cablemaker as been doing since the ending of the artificial conditions of var time. I propose to outline first some of the more recent levelopments in materials and processes and then to show the mpact these have had on cable design and construction.

Materials and Processes

As a conductor, copper remains virtually unchallenged and, xcept for features such as compacted strands and conductor creening, substantially unchanged for general cable use, although or certain high-temperature applications conductors are electro-plated with nickel or silver.

Aluminium, with its relative lightness, low cost and reasonable conductivity, is at first sight attractive, and certain of its uses are vell established. Technically, its wider application presents no lifficulties, and cablemakers are ready to employ it. Its ultimate cale of use depends on relative costs, which must include the greater volumes of insulating and protective materials required or aluminium conductors compared with those for the copper equivalent.

With rubber, the main advances have been in manufacturing echniques, such as improvement in extrusion machines and the ncreasing adoption of continuous vulcanization. Conducting rubber is being explored, in place of metallic wires, for the screens of colliery trailing cables; and rubber-insulated cables are protected from ozone and oil by a thin external layer of synthetic rubber.

Of the many synthetic materials developed abroad, only poly-chloroprene is in extensive use; its properties make it valuable as a cable-sheathing and protective covering in aircraft, mining and other spheres. Silicone rubber, with its high heat resistance and flexibility at low temperatures, is employed in aircraft wiring.

The war-time developments of plastics have been pursued further, giving the cablemaker a wide range of materials not available in any substantial form in pre-war days.

Polyvinylchloride (p.v.c.) and its co-polymers give a variety of insulating and sheathing compounds, easy to process and characterized by toughness, flexibility and resistance to many solvents and to burning.

Polyethylene, available in several grades with differing physical properties, has low moisture absorption and outstanding electrical characteristics for high voltages and high frequencies. It calls for special precautions in processing and use to counter the effects of exposure to heat, to sunlight and to contact with a variety of other materials.

Polystyrene, also with low moisture absorption and excellent electrical properties, has a tendency to brittleness, which is overcome by using orientated tapes to form laminated dielectrics.

Polytetrafluoroethylene (p.t.f.e.), chemically and physically stable, has many of the virtues of polyethylene and unique qualities for high temperatures up to 250°C, such as for aircraft wiring. It is costly. Fabrication demands special processing techniques, and high-temperature uses require conductors of nickel, silver or plated copper.

Nylon has toughness which renders it valuable for sheathing, extruded or as a textile braid.

In papers and compounds for mains cables, the significant developments since pre-war days have been in the quality and choice of materials.

Improved wood-fibre papers have largely replaced the earlier-favoured manila and manila-wood papers. The petroleum refiners have made available much-improved grades of hydro-carbon oils. Established precautions against migration of fluid compounds on gradients are now supplemented by a method of mass-impregnation employing an impregnant which, though fluid at cable manufacturing temperatures, is a plastic solid up to temperatures higher than those of service.

The post-war scarcity of lead stimulated the search for alter-native materials of lower density and greater mechanical strength.

The lightness, strength and high fatigue resistance of alu-minium make it an obvious choice for cable sheathing. It makes armouring unnecessary for most uses and obviates the need for reinforcement in pressurized cables. Satisfactory corrosion protection has been evolved, and jointing techniques are well established. In this country, most experience is with sheaths made by the "tube-sinking" process of pulling the cores into an oversize tube which is subsequently drawn down to the desired diameter. Alternatively an over-size sheath may be formed from longitudinal strip, argon-arc welded, the sheath then being drawn down to size and helically corrugated to improve its flexibility. Direct extrusion, although attractive, presents diffi-culties because of the high temperatures and pressures involved. Much work has been done in this development, both here and on the Continent. One British cablemaker has installed an extrusion press, using solid billets and operating at a reasonably low temperature, which gives promise of the satisfactory pro-duction of aluminium-sheathed cables in long continuous lengths.

A number of composite-sheath constructions have been evolved for telephone cables, to provide a moisture seal and resistance to corrosion. An example is the American sheath, embodying aluminium and steel strip and covered with extruded polyethylene. Thin lead sheaths with an outer layer of polyethylene are expected to find increasing use.

A significant development in protection materials is the replace-ment of impregnated textile servings by tapes of self-sealing rubber and p.v.c. to reduce the danger of water absorption.

Impacts on Cable Design and Construction

Recent developments in telecommunication cables have made full use of plastics, leading not only to the replacement of estab-lished types of insulant on technical or economic grounds, but also to the evolution of new types of cable with very exacting

Mr. Dunn is with British Insulated Callender's Cables, Ltd.

transmission characteristics, for telecommunication, radio and electronic equipment.

In switchboard wiring, coloured and striped p.v.c. is tending to replace the widely-used enamel and textile coverings.

It is significant that in the first 40 years of this century, frequencies to be transmitted by cable advanced a thousandfold from 3×10^4 to 3×10^7 c/s, and that during the succeeding decade they rose to 3×10^{10} c/s, a further thousandfold increase. Only polyethylene and polystyrene, with their low loss factors, could meet the requirements of long-distance and high-frequency transmission. Even these materials would have been inadequate had it not been for the concurrent development of new techniques to give precise dimensions and of new electronic test-gear.

Polyethylene, the more easily handled material, is widely used. As a solid extrusion it is being introduced for telephone distribution cables. Radio-relay and television-relay cables have polyethylene insulation and sheaths; they are light in weight, and have low transmission losses and high moisture resistance. Polyethylene insulation is being employed for quad-type carrier telephone cables; on the Continent, polystyrene string and tape constructions have been adopted for this purpose; and experiments are being made in America with cellular polyethylene. The conventional television down-lead cables are polyethylene insulated: for the higher frequencies in Band III of the coming commercial programmes an expanded polyethylene insulant has been introduced.

A notable use of plastics is in the spaced-disc coaxial pairs developed for multi-channel telephony or television signals. A classic example is the 1 in coaxial cable forming the London–Birmingham television link, and many hundreds of miles of $\frac{3}{8}$ in coaxial pairs are in regular use, both at home and abroad, operating on frequencies up to 10 Mc/s and providing up to 600 speech channels.

The recently-announced transatlantic telephone cable is the first project of such magnitude to be undertaken anywhere in the world: it is to have solid polyethylene insulation of thickness and precise dimensions which the recent advances in technique have alone made feasible.

Polystyrene in the form of multiple tapes applied in an open helix is used for several cables of coaxial type up to $3\frac{1}{8}$ in diameter, employed for high-power high-frequency links. There are numerous radio-frequency cables with solid or semi-air-spaced polyethylene insulation.

Television cameras require multiple connections for video frequencies up to many megacycles per second, for audio signals and for power and control circuits. These complex cables must be small, flexible, rugged and weatherproof, and must be easily capable of extension. Accurately extruded thin polyethylene coverings on small conductors, and a special connector moulded to the cable, have met these requirements.

For lighting and low-power cables, increasing use is made of rubber-insulated tough-rubber-sheathed cable. Advances in technique have permitted reductions in diameter of many types of cable, as reflected in B.S. 7: 1953. The use of p.v.c. wiring cables is growing, with emphasis on conduit wiring. B.S. 1557, for polyethylene-insulated and p.v.c.-sheathed types, has just been revised.

Aircraft cables provide a good illustration of the exploitation of the special properties of plastics, in that light weight, resistance to solvents and a wide temperature range of operation are essential.

The insulation of pre-war aircraft cables was rubber, protected by a lacquered cotton braid. This has in late years been replaced by glass-fibre braid and a polychloroprene sheath, with a saving of 50% in weight and a wider temperature range.

For the extreme temperatures encountered in jet aircraf combinations of p.t.f.e. and glass-fibre braid, though costly, giv outstanding performance, and the pre-war operating range o $-10°$ to $+70°$C has thus been extended to $-75°$ to $+400°$C.

Broadly speaking, paper-insulated cables continue to hold th field for power supply. Cables with plastic insulation and sheat have proved satisfactory in performance in the lower-voltag range. They are attractive because of light weight and simplicit of jointing. In countries where skilled jointers are available any marked extension in their use seems likely to turn on relativ costs.

In the very-high-voltage field, the post-war years have seen notable extension in the use of higher voltages, for which ther has been an accentuated swing to the use of pressurized cables The higher operating stresses and temperatures and increase circuit capacity of these cables are reflected in the lower cos of an installation; and this, coupled with the fact that they ar essential for voltages over 66 kV, is responsible for the greatl increased demand. Pressurized cables are of four main types namely:

Oil-filled cables, which are impregnated with low-viscosit oil and embody channels to permit longitudinal oil-flow to o from reservoirs. The static operating pressure is normall 75 lb/in^2. A later variant, in which the three cores are lai side by side within a flat-sided sheath, enables reservoirs to b dispensed with, the cable being "self-compensating."

Gas pressure cables, which are subject to internal gas pressur of 200 lb/in^2, and are of two sorts. In the *gas-filled cable*, th dielectric is formed from pre-impregnated paper tapes, the ga filling the interstices between tapes to form a composite dielectric *Impregnated pressure cables* are fully impregnated, the gas bein accommodated in under-sheath clearances.

Compression cables, which consist essentially of solid-typ cables covered by a thin lead diaphragm and enclosed in a outer sheath or pipe. The space between diaphragm and pip is charged with gas, the pressure of which is transmitted to th cables via the diaphragm.

Pipe cables, which are oil-impregnated solid-type cable without sheaths, and are drawn into a steel pipe containing oi or gas under pressure.

Although 220-kV oil-filled cables were installed in Paris i 1936, there has in the past been no general call in this countr for cables for voltages above 132 kV. However, since th 275 kV system now under erection will eventually require cable for that voltage, the B.E.A. and cable manufacturers hav collaborated in the installation of trial lengths of 275 k pressurized cables.

Cables for even higher voltages have been made for tw installations abroad. For the Kemano project in Canada, tw runs of 301 kV oil-filled cable have been laid, one having alu minium conductors and sheaths; while in Sweden a 380 kV cabl of the oil-filled type has been installed.

Another post-war trend is towards the use of high-voltag submarine crossings. A project has been worked out for 132 kV cable, crossing the Channel, to link the British an French systems, and sea trials have been made, using gas-fille and compression cables.

In Canada, a British cablemaker is shortly to install an 18-mil underwater link, between British Columbia and Vancouve Island, consisting of 138 kV gas-filled cables.

This is the picture as I see it. It may not be comprehensive o balanced, but I hope I have given an impression of the cable maker's activities. The work continues; and I have no doub some future Chairman, choosing the same topic, would fin ample new material for his Address.

Abstract No. 1819
Feb. 1955

NORTH-WESTERN CENTRE: CHAIRMAN'S ADDRESS

By Professor E. BRADSHAW, M.B.E., M.Sc.Tech., Ph.D., Member.

"LABORATORY WORK IN ELECTRICAL ENGINEERING COURSES"

(ABSTRACT *of Address delivered in* MANCHESTER, *5th October*, 1954.)

This Address relates to laboratory work in general, but special attention is directed to the problems of the initial stages.

Aims and their Realization

Laboratory work should aim at:

Direct tuition in the standard methods of measuring electrical quantities and the use of standard items of measuring equipment.

Understanding of topics treated theoretically in lectures and appreciation of the limitations of the idealized treatments which are of necessity used in much lecture work.

Familiarization with the appearance, construction and capabilities of apparatus, and with the magnitudes of electrical quantities.

Cultivation of a critical attitude towards all observations, extending to fields other than electrical. As experience is gained the student should acquire a feeling for the organization of his experimental work and, in the larger experiments, some sense of co-operative activity.

Development of the writing of concise English and the presentation of data in a clear and understandable form.

The successful achievement of these aims depends on:

Selection and Design of Experiments.

These are discussed more fully below.

Laboratory Instruction Sheets.

These, together with the personal attitude of supervisors, are responsible for developing a critical attitude towards the experiments. They should be as comprehensive as possible, and should include a number of questions to be answered to ensure that the implications of the experiment are appreciated, and that the relative importance of errors in the different readings are noted; together with additional notes explaining why particular items of equipment are used, and linking the equipment with others or with future course work. In the early part of the course, detailed instructions are only too necessary for most students to formulate a reasonable approach and presentation. In the later years the instructions should be curtailed.

Report Writing

The traditional procedure is for the student to take away his experimental "results" after taking whatever steps seem necessary to apply checks on their completeness and reasonableness before leaving the laboratory, and then to submit after suitable private work, a finished report. The method has the disadvantage that the student has to do much of his work away from the advice and guidance which he may need in interpreting his results.

Ideally, reports which have been written in the laboratory would seem to meet these objections. Certain experiments, however, require literature to be consulted and lengthy calculations made. In these circumstances, the best compromise appears to be for the student to do as much as is practicable before leaving the laboratory. This laboratory record, duly checked and initialled by the instructor before the student leaves the laboratory, is then submitted, together with any additional work which is done later, as a final report.

It is worth impressing upon students the degree to which a carefully annotated diagram can express concisely the experimental conditions and procedure.

Laboratory Organization

A typical elementary electrotechnics laboratory may cater for 15–25 groups of students. With the numbers involved, it would be impossible, even if it were desirable, to allow the students to choose their own meters, etc., and formulate their own procedure. This practice may have advantages where small numbers are concerned and generous supervision is possible, i.e. in later stages of the course, but it leads to unfinished experiments, unchecked errors and too much unproductive blundering during the early stages where students are unfamiliar with equipment.

Supervision

Supervision of elementary laboratories is difficult as it tends to be an unpopular staff duty, which then devolves upon young demonstrators. This may appear to solve the staffing problem, but such supervisors do not always take the critical viewpoint which is necessary, they change frequently so that there is no continuity, and they may have no responsibility for assessing the students. Senior staff who take such duties seriously give a valuable service.

"Original Work"

Little of this is possible during the first two years of the course. To encourage students during this period to think of the type of equipment to choose, and the method of approach to problems, they can be asked to design experiments, on paper, to illustrate certain topics or verify experimental laws, specifying fully the type of meters and other equipment required and the handling of the results.

Selection of Experiments

One of the problems throughout the course, but particularly at the start of the first year, is that of students undertaking laboratory experiments before the relevant theory has been covered in lectures. A lecture may be easier to assimilate and possibly more interesting to a student who has encountered the problem in practice, but many experiments cannot be performed properly without a knowledge of the theoretical background, and that cannot always be condensed sufficiently to appear on an instruction sheet.

This difficulty can largely be overcome, and the interest in what may be termed "experimental method" aroused, by commencing the laboratory work of the initial year with experiments requiring only an elementary knowledge of d.c. circuits, but requiring observations on the accuracy of the meters used, the scatter of observations, linear and non-linear relationships, curve-fitting and the use of logarithmic graph paper, and simple energy considerations. These may occupy the first ten weeks whilst the lectures are covering fundamental theory. A fuller treatment of experimental method, considered as the study of the factors governing the design of experiments and the analysis of experimental results, is appropriate to a later stage of the course, in the measurements laboratory.

Subsequent to the special arrangements for the first few weeks, the choice of experiments is governed by (*a*) *suitability*—simple experiments calling for little ability from the students lead to loss of interest and should be avoided. Such experiments, however, have their uses when emphasis is to be put upon the numerical analysis of results; (*b*) *practicability*—experiments must be so arranged that, in competent hands at least, they perform satisfactorily, and it is then possible to point to incompetent hands, and not unsatisfactory equipment, as the source of error.

There is no doubt that greater use should be made of demonstrations both in the lecture room and in the laboratory, as a supplement to formal experimental work. Besides the stimulation which such demonstrations provide, the student can be introduced to a number of topics or illustrations which would not merit treatment as formal experiments.

Prof. Bradshaw is Professor of Electrical Engineering in the Faculty of Technology, University of Manchester.

A certain amount of time devoted to examination of and reports on equipment, fault location, elucidation of circuits and the production of schematics from such experiments, is not out of place.

It is commonplace that the field of study included in Electrical Engineering has been and is expanding at a remarkable rate. The increased range and content of the science can be met to some degree in lecture courses, by placing increasing emphasis on systematic and generalized treatments. There is a counterpart to this action, in laboratory work, in the introduction of experiments to illustrate and investigate generalized topics such as, for example, circuit theorems, field distributions, and elements of control systems, etc.

In the face of an ever-increasing volume of material calling for inclusion, traditional laboratory experiments must be constantly examined for justification. In the earlier days of the science when the subject was less extensive, experiments were performed which to-day may be found to be either redundant or lacking in basic scientific value. The older branches of the science, e.g. electrical machinery, are often the most in need of such attention.

Apart from the demand of new material for inclusion, the possibility should always be borne in mind that existing and basic material may not be adequately treated. In spite of the large proportion of laboratory time given in many courses to electrical machinery, it is probable that such basic phenomena as commutation, armature reaction and the idea of the machine as an amplifier are often inadequately treated experimentally in the early stages.

Design of Experiments

Very little apparatus, except for standard measuring gear, is commercially available in the most suitable form for electrical engineering laboratories. There is almost limitless scope for ingenuity in the design of good experimental layouts.

Certain standard electrical machines are more or less suitable in their commercial form, but even in this field the special requirements of an educational institution often call for special designs. Such requirements include electrical ratings to conform to laboratory supplies, special characteristics, the addition of search coils and thermocouples, and facilities for driving or loading.

Even when no modification of an existing experimental equipment seems necessary, it will often be found that the experimental procedure can be revised with advantage. Students should not be asked, for example, to repeat a series of operations at a number of different loads, speeds, voltages, etc., unless some really significant result is thereby illustrated.

Electrical machinery and a few other applications excepted, one usually has to begin a design of an experiment *ab initio*.

Briefly, equipment for use in electrical engineering instructional laboratories should as far as possible possess the following features:

- Clear exposition of the essential principle it is designed to elucidate.
- Educational value in stimulating insight and experimental facility.
- Time-saving: not burdened with non-essential manipulative difficulties.
- Modern in technique and design.
- Good quality, reasonably "student-proof," and capable of proper maintenance.

Simplicity is important, particularly in the elementary stages. Students should not be confused by an unnecessarily bewildering
array of apparatus whose function is far from clear. Cle
layouts, ample labelling of apparatus in respect of rating
polarities, etc., are essential. It is important to subordina
auxiliaries so that they do not overshadow the essential apparatu
in complexity, operational difficulty or size. Suitable relations ar
perhaps unconsciously, fulfilled in machine and large-scale pla
tests, but they become more and more difficult to realize in benc
experiments.

Whilst the exercise of connecting up apparatus is valuable,
can assume too great an importance. There are certain exper
ments where a clear mental view of the essential circuit throughou
the experiment justifies the use of permanently connected equi
ment, particularly if this can be associated with a semi-pictori
layout. The use of models, besides illustrating a much-use
engineering technique, often permits the justifiable exaggeratio
of certain significant phenomena. How far a model techniqu
may be justified if secondary phenomena are thereby modifie
is a matter for careful consideration.

The application of the more recent measurement techniqu
to older branches of the subject should be considered. Electric
machinery laboratories have tended to be inhabited only b
voltmeters, ammeters and wattmeters. There are signs, how
ever, that electronic devices and a.c. bridges and potentiomete
are finding useful application here. Apart from their value a
instruments, it is good for students to see the inter-relation of th
various branches of their studies.

Laboratory Design

Surroundings which are not only technically suitable but als
attractive and workmanlike in appearance assist in providing th
right atmosphere for the efficient performance of experiment
work. The opportunity to design the laboratory building itse
seldom arises. When it does, the ideal arrangement appears t
be a room with as much unbroken lower wall space as possibl

Arrangements for electrical supplies to the work benches shoul
be flexible but not extravagantly elaborate. The wiring shoul
however, be planned so that no additions in the form of unsightl
temporary wires are needed as far as can be foreseen.

Benches should provide adequate storage space for portab
equipment in cupboards and drawers. It is desirable to have
certain proportion of plain benches under which wheele
apparatus can be placed, or with a single low shelf for th
storage of the larger electronic test units. It may be possib
to arrange for tables carrying permanently connected assemblie
e.g. potentiometers, to have false tops to permit normal use o
the tables for other purposes.

Conclusion

Whatever the material facilities provided for the student, h
effective development as an electrical engineer, particularly i
the experimental field, will be determined largely by the enthu
siasm of the staff concerned. As has been mentioned earlie
the work of those members of the staff who take seriously th
design, development and operation of laboratory facilities is o
the greatest value. Such work should command the respe
which it deserves and should not suffer in academic regar
relative to other activities. Not only does it contribute im
measurably to the education of the student, but it forms on
of the best channels through which assessment of the student
abilities may be made. In general, insufficient use is made o
such opportunities for assessment.

Abstract No. 1850
Mar. 1955

NORTHERN IRELAND CENTRE: CHAIRMAN'S ADDRESS

By Major P. L. BARKER, B.Sc., Member.

"RECENT TELECOMMUNICATIONS DEVELOPMENTS IN NORTHERN IRELAND"

(ABSTRACT *of Address delivered at* BELFAST, 12*th October*, 1954.)

The Belfast–Dublin Coaxial Cable

The Belfast–Dublin coaxial cable was laid at the beginning of the present decade and was partially brought into use at the beginning of 1953 by using the audio pairs provided for intermediate circuits; it was brought into full use in September, 1954.

The cable consists of two tubes, each having an inside diameter of 0·375in and containing a central solid-copper conductor having a diameter of 0·104in and located by slotted polythene discs at 1·3in intervals. The outer tube is a copper strip or tape with crimped edges for interlocking when formed into a tube. Each tube is reinforced with two steel tapes laid together with packing to form a cylinder, and the whole unit is wrapped with paper tape. The coaxial element of the cable includes eight 20lb unloaded pairs for control purposes and two 40lb screened pairs with 16mH loading for music transmission. Between Belfast and Newry the cable also contains a varying number of 20lb conductors for audio circuits.

The frequency range is 60–2 604kc/s, and within this range it is possible to accommodate 600 speech channels, each covering the band 300–3 400c/s. A separate tube is used for each direction of transmission, and it is necessary to have repeaters at approximately 6-mile spacing. There are nine such repeater stations between Belfast and the Border, and ten between the Border and Dublin.

The method adopted for locating the speech channels over the frequency spectrum is as follows. Speech channels are divided into groups of 12, and each channel in a group is modulated with a different frequency so that the whole group occupies a band of 60–108kc/s.

The next process is to take five groups and to modulate each of these with a different frequency, the whole forming a supergroup in the band 312–552kc/s. The system can be set up using one or more supergroups up to a maximum of 10, giving a range of channels from 60 to 600. The bandwidth of each supergroup is 240kc/s, with a space of 12kc/s between supergroups 1–2 and 2–3 and 8kc/s between other supergroups.

The system also requires the provision of two pilot carriers at 60kc/s and 2 604kc/s to control the gain and frequency response of the line amplifiers in the two directions of transmission, and to provide comparison between the carrier frequency generation equipment at the two terminal stations.

The equipment at each station operates on a 230-volt single-phase 50c/s supply, and for the intermediate repeater stations the supply is fed over the coaxial tubes from Belfast and Newry, the load at each station being about 250 watts. Standby emergency power plant is provided at the terminal stations for use in the event of a mains failure.

To meet the initial demand for circuits between Belfast and Dublin the route has been equipped with one supergroup of 60 circuits using supergroup 3, which lies in the band 564–804kc/s. This can be added to in the course of time as traffic increases. No carrier circuits are used or envisaged between intermediate towns in Northern Ireland.

Three types of repeater are used, namely terminal repeaters, dependent repeaters and power-feeding repeaters. A stabilized power supply at 650 volts is fed to the central conductors of the two coaxial tubes via a power filter, and is fed to the line through low-pass filters; high-pass filters in series with the carrier equipment prevent this 50c/s supply from entering the repeater equipment. At each dependent repeater station the outgoing supply is stepped up to the line voltage by means of an auto-transformer before being applied to the next section of the line.

The terminal repeater amplifies the output from the supergroup translating equipment, which is at a level of -45dB, to the line level of -15dB. The amplifier is a three-stage unit with a fixed nominal gain of 35dB and a flat response over the entire frequency range. The two line pilots of 60kc/s and 2 604kc/s are added at this point, and also the mains supply for feeding the dependent repeaters.

The pilot frequencies are generated at the terminal station, and the amplitude of one or both of these is used at each repeater station to control the gain of the line amplifier. The pilot level is stabilized at each terminal station to a level of -25 ± 1dB into 75 ohms at the output of the transmit amplifier.

The method of gain control is as follows. The pilot frequencies are separated amplified and rectified, and the resulting direct potential is used to bias two 2kc/s oscillators so that the amplitude of oscillation varies according to the level of the appropriate pilot. These 2kc/s tones are fed to thermistors in the negative-feedback path of the line amplifier so that the variations in the 2kc/s tone derived from the 60kc/s pilot alter the amplifier characteristic at the low-frequency end of its range, whilst variations in the 2kc/s tone derived from the 2 604kc/s pilot alter the line amplifier characteristic at the high-frequency end of the band.

The 60kc/s pilot control is not used at the intermediate dependent repeater stations, as changes in cable attenuation are less at the lower frequencies than at the high, so that the levels will remain closer to the predetermined working levels for longer distances at the lower end of the frequency range.

Four wires of the intersticial pairs in the cable are used to indicate faults at any of the stations, the circuits being completed by an earth return. The four types of failure for which alarms operate are valve ageing, h.t. failure, and variation of the levels of the two pilot frequencies by more than 5dB. The valve-ageing alarm is classed as non-urgent, and the three others as urgent and requiring immediate attention.

Telegraph Services

The installation of the coaxial cable, with its audio pairs, enabled a more reliable telegraph system to be provided between Belfast and Dublin, and a 12-channel v.f. telegraph system was installed at the same time, using two of the intersticial pairs.

The automatic switching system was fully introduced into Northern Ireland in February, 1954. This system is based on

Major Barker is with the General Post Office.

the principles employed in automatic telephony and uses the same type of equipment for switching a teleprinter circuit from any office to any other in the United Kingdom without the intervention of an intermediate operator.

The layout of the trunk network is based on the provision of 22 switching centres, which are divided into three categories called zone centres, switching centres class I and switching centres class II. There are six zone centres which are all directly connected with each other and to some switching centres of each class; there are eleven switching centres class I, which are connected to some zone centres and to certain of the other switching centres; and there are five switching centres class II, each of which is directly connected to at least one zone switching centre and to certain other switching centres depending upon the volume of traffic.

The numbering system is based on mixed letter and figure code and the areas served by each switching centre. By this means it is possible to issue a common directory for the whole country. From this an operator at Coleraine with a message for Newark would find from the directory the dialling code for Newark given as "NG 221." He would then find from the switching-centre dialling-code list that the number for Nottingham is "325," and so he would dial "325–221" and get straight through to the teleprinter at Newark and then immediately proceed to transmit his message. There are facilities provided for answer back from the distant end, busy signals, out-of-order signals and various alarms.

Television Service to Belfast

In the autumn of 1952 it was decided to set up a temporary television station in Northern Ireland to serve the Belfast area, and public transmission was to be available in time for the Coronation. The problem facing the Post Office in this enterprise was to provide a link between some point on the existing television network in Great Britain and the new station. Normally this would have been provided by means of a cable or radio link from the nearest point (Kirk o'Shotts), but to do this in the time available, including the manufacture of all the equipment, would have been impossible.

Two alternative methods were considered: to receive on the Ayrshire coast the signals radiated from Kirk o'Shotts and relay them across the Irish Sea on a 200 Mc/s link for which equipment was available; or to receive the direct transmission from Kirk o'Shotts at a suitable site near the proposed B.B.C. transmitter. Measurements of the Kirk o'Shotts field strength made both on the Ayrshire coast and near Belfast showed that the relay scheme had no advantage over direct pick-up near Belfast, which would also be a much simpler scheme to get going in the time, and it was therefore adopted.

The site chosen for the Post Office receiving station was situated at the top of the Black Mountain at about 1 200 ft, and the B.B.C. temporary transmitter was to be located at Glencairn at about 400 ft and a mile away from the Post Office receiving station. The link between these two sites is by means of a coaxial cable. This is a standard two-tube cable with 16 interstician pairs which are used for supervisory and control purposes. The cable is armoured and is laid directly in the ground. Although the route is short it is extremely hazardous, having gradients greater than 1 in 2 in parts, and there is also a difficult ravine to negotiate. Here the cable has been suspended from stranded steel wires slung between poles on the banks.

The field strength measurements of Kirk o'Shotts at the summit of the Black Mountain showed that normally about $100\,\mu$V/m would be received and that an aerial array having a gain of some 15 dB over a half-wave dipole would be necessary to give a margin of about 25 dB against fading before the signal became unusable.

The aerial array consists of twelve inverted-V elements slung from a triatic supported by three 45 ft poles, the ends of each V being made off on to short poles. The twelve elements are combined into two groups of six to provide two entirely independent arrays each having a theoretical gain of 16·5 dB. The measured gain of each array was found to be 15 dB. The outputs from the arrays are matched to a 75-ohm coaxial cable for leading to the receiver.

This type of aerial system has the disadvantage of having a very broad bandwidth and will pick up signals down to the medium-frequency band, but the most serious source of interference against which protection had to be provided was the adjacent B.B.C. television transmitter on 45 and 41·5 Mc/s. To help overcome this trouble a high-pass filter is included between the array and the receiver, and this gives attenuations of 28 dB at 45 Mc/s and 40 dB at 41·5 Mc/s. Rejection circuits are also included in the amplifier following the filter.

The output from the high-pass filter is passed through a pre-amplifier to the receiver, which is fitted with automatic gain control. The television modulated carrier cannot be used for this purpose because its amplitude varies with the brightness of the transmitted picture as well as the signal strength of the carrier. Following each synchronizing pulse in the television signal there is a short period of black level during which the carrier amplitude represents the signal strength and can be used for control purposes. This particular portion of the signal is used to produce a direct voltage which keeps the black level of the receiver output constant. The receiver output is a video signal in the 0–3 Mc/s band with a peak-to-peak amplitude of 1 volt.

For transmission over the coaxial cable to Glencairn, a carrier of 6·12 Mc/s is used, modulated by the video signal from the receiver. At Glencairn, after equalization, the signals are demodulated, amplified and fed to the B.B.C. transmitter through short coaxial leads.

The power supply for the equipment is from the public supply mains at Glencairn. For the equipment at the top of the mountain the power is fed over the coaxial cable from Glencairn, the radio equipment being protected by power filters. Two independent feeds are provided, one on each tube, and up to 10 amp can be fed by this means, which is adequate for the equipment and also provides for lighting, heating and test equipment. In addition, an emergency 5 kW Diesel set has been provided at the top of the mountain for any necessary additional heating or cooking facilities required, and it can also be used to run the equipment in the event of faults which require the cable to be isolated from the mains supply.

The television link to Belfast has now been in service for about 18 months, and since the scheme is of a somewhat experimental nature, it will be of interest to see how the link has behaved in practice.

The 125-mile path from Kirk o'Shotts to Belfast is, of course, non-optical, and it was anticipated that there would be severe fades in the signal to an extent which would make it unusable. Their magnitude and duration were an unknown factor. Experience to date has shown that deep fades do occur, but so far they have not seriously interfered with the service. The amount of lost time has averaged about 0·25% of the total transmission time, or about 8 min a week.

The behaviour of the equipment itself has been excellent, and the total amount of lost time from this cause, since the service started, has been less than 2 min.

Abstract No. 1822
Feb. 1955

NORTH STAFFORDSHIRE SUB-CENTRE: CHAIRMAN'S ADDRESS

By J. M. FERGUSON, B.Sc.(Eng.), Member.

"SOME ASPECTS OF HIGH POWER PULSE GENERATORS"

(ABSTRACT *of Address delivered at* STAFFORD, 11*th October*, 1954.)

The general arrangement of an impulse generator, shown in Fig. 1, consists of the following units: a *supply*, capable of delivering all the energy required by the pulse generator as a

Fig. 1.—Basic scheme for pulse generator.

steady load; a *store*, into which the supply feeds, and taking one of several forms discussed later; a *convertor*—the energy stored need not be electrical, and as high-energy electrical impulses are being considered the store is coupled to the load by a convertor to convert the energy to an electrical form; a *switch* which controls the discharge into the load at the appropriate instant; if the whole energy is not to be used for each pulse, the switching device must be capable both of a making and a breaking duty.

Energy Store.—Any type of energy store can be used, provided it is capable of repetitive cycles.

Five practical types of store which are used in a number of pulse generators will be examined: potential mechanical energy, dynamic energy, electromagnetic energy, electrostatic energy, and electrochemical energy.

considerable losses which could be reduced by running either in a vacuum or in hydrogen, but the increased complications are not often considered worth while. For short pulse lengths, the only satisfactory store is electrostatic.

Switch.—Conventional contactors for circuit-breakers exist to meet most conditions where low repetition and long pulse rates are involved. Air-break and air-blast types are the most likely to withstand heavy duty.

Special consideration has been given to making and breaking devices which has enabled mechanical switches to be developed that will carry out switching operations at repetition rates of 50 times per second with capacities of several thousands of amperes. These switches are used in a mechanical rectifier.

For high repetition rates, high-vacuum triodes can be used, but the maximum voltage is limited by the emission of the filament at the peak voltage that the valve will withstand.

Gas discharge devices of the following different types have been used: hot-cathode gas discharge valves, spark-gaps, cold-cathode diodes, and mercury-pool valves. The appropriate characteristics of each of these types of switch are discussed.

A further type of switch consisting of a saturable iron-cored reactor can be used. High powers with short pulse lengths have been reached with this type of device.

Table 1

Type of energy store	Volume efficiency		Duration of pulse discharge		Storage losses	Remarks
			Max.	Min.		
	ft-lb/ft³	joules/cm²			watts/ft-lb	
Potential mechanical	$3\cdot12 \times 10^3$	0·15	Hours	Hours	Nil	Based on water storage
Dynamic	$5\cdot0 \times 10^5$	24	Seconds	Milliseconds	$6\cdot0 \times 10^{-4}$	—
Electromagnetic ..	250	0·012	Milliseconds	Milliseconds	30	Losses based on 10 millisec
Electrostatic ..	$1\cdot15 \times 10^3$ 20·5	0·052 0·001	Hours	Microseconds	Very low	Impregnated paper condensers
Electrochemical ..	$1\cdot6 \times 10^6$	78·5	Hours	Seconds	Very low	Planté cells

The following characteristics of these types of store, summarized in Table 1, are considered: space required for a given unit of energy stored, suitability for particular pulse length, and storage losses.

Of these types of store the most efficient for its size is the electrochemical store, but unfortunately, owing to high internal impedance and polarization effects, short-time currents are severely limited. The 3-sec rating is normally not greater than 22 times the 10-hour rate of current discharge, although high-duty cells can be obtained in which this is very significantly increased.

For medium pulse lengths, the most generally used store is a rotating mass such as a flywheel or rotor. This gives rise to

High-Power Pulse Generators.—Pulse generators can be used in the following general ways:

Where the load fluctuates rapidly, as in rolling mills.

Where conditions caused by the release of a single high-power pulse of energy for a short time are to be investigated, e.g. particle acceleration, cloud chambers, short-circuit testing and high-voltage impulse testing.

Utilization of echo techniques for producing a high-power burst of radiation.

The design of one example of each type of generator is considered in relation to the fundamental characteristics outlined in the Address. For this purpose, the design of a hot reversing mill, a short-circuit testing plant and a pulse generator for radar, using a mercury-pool modulator valve, are considered.

Mr. Ferguson is with the English Electric Co., Ltd.

Abstract No. 1823
Feb. 1955

RUGBY SUB-CENTRE: CHAIRMAN'S ADDRESS

By D. EDMUNDSON, B.Sc.(Eng.), Member.

"THE CHANGING OUTLOOK IN ELECTRICAL MEASUREMENTS"

(ABSTRACT *of Address delivered at* RUGBY, *5th October*, 1954.)

A singular feature of present-day electrical engineering is the increasing emphasis on mathematics, not only as a necessary basis, but as a practical tool, for the solution of specific problems. Fifty years ago, when electrical engineering in Rugby was in its infancy, the mathematical basis of electrical machine design was being worked out by such men as Dr. Carter, who produced a complete solution of air-gap field problems, and Field and Taylor, who worked out solutions of the problem of eddy-currents in stator windings. In those days, experimental techniques were inadequate and many of the components on which we now depend for measurement work did not even exist; but to-day the position is entirely changed. It is difficult to point to an example of a measurement problem in electrical engineering which is incapable of solution in a practical, convenient and reliable way.

In the belief that this striking change in the possibility of measuring quantities only capable of mathematical estimation in the days of the pioneers has not been realized by many engineers, three simple examples were illustrated, with demonstrations.

Measurement of Torque.—In the year 1906, long before the introduction of conventional torque-meters, an interesting proposal was made by Mr. A. P. Young. He pointed out that the instantaneous value of the rate of deceleration of the rotor of a machine—a quantity required in the course of testing—could be obtained directly in a simple manner. A d.c. generator driven by the shaft could produce a voltage proportional to speed, so that the rate of change of this voltage would be proportional to that of angular velocity, and hence to torque. This quantity Young proposed to obtain either by applying the voltage to a condenser, whose charging current would be its derivative, or by deriving from it a current to be passed through the primary winding of a mutual inductance; the secondary voltage would then be proportional to deceleration. Young actually attempted the method, but was unsuccessful only because instruments of the necessary sensitivity did not exist. Many years later, Dannatt and Redfearn used the same principle to measure the accelerating torque of an induction motor. Having now available both the valve amplifier and the Duddell oscillograph, they produced highly successful torque/time curves. Finally, a demonstration was given of the torque/speed curve which can be obtained using a cathode-ray oscillograph. Thus, a complete solution of what was an intractable problem 50 years ago has been obtained by modern measuring techniques.

Measurement of Angular Velocity.—Although for many years reliable tachometers have been available, these have had two main drawbacks. None has been capable of high precision, and all have imposed a load on the driven shaft which is unacceptable for light-power motors. The first condition was met during the late war by adapting the principle of Maxwell's commutator bridge, which was demonstrated. Long neglected because the apparent complexity of the traditional expressions for balance were considered to render it approximate only, the method is capable of high accuracy in an indicating instrument. This was achieved using components which had been available for many years; but its practical realization as, effectively, a pulse-counting device made its direct application to various forms of drive a simple matter. Using a magnetic pick-up and amplifier, speed can now be measured provided only that a shaft has some magnetic irregularity; failing even this, a photo-electric cell and amplifier can convey to the tachometer pulses derived from black and white markings. It is therefore easy to measure angular velocity with the highest precision without imposing any load whatever on a rotating shaft—another achievement which was out of the question 50 years ago.

Measurement of Temperature.—As a final example, the use of thermocouples presents a further interesting demonstration of the change brought about by modern techniques. Until recently, their use at moderate temperatures was restricted to permanently installed, carefully engineered installations, or delicate null-methods, by the small value of the e.m.f. produced. This was insufficient to operate robust, portable instruments. A development of the potentiometer led to the negative-feedback chopper amplifier. In this, d.c. amplification is achieved by converting the d.c. signal to a.c. by mechanical means and reconverting it after amplification. The function of the amplifier is then to compare the e.m.f. from the thermocouple with that produced by the amplifier output current in a fixed resistor, and to ensure that there is negligible difference between them. This instrument is now portable, robust and accurate; indications are obtained as quickly as a d.c. ammeter can be read, while the resistance of the leads has no effect on the indications. Temperatures can be obtained from a variety of points in quick succession, and couple wires can be so thin that they do not abstract heat from materials of poor conductivity.

These examples, few as they are among a multitude which could have been chosen, may serve to illustrate the difference between the approach to practical problems which is now open to the electrical engineer and that to which his predecessors were confined. How is he to take advantage of the new measuring techniques? Although measurement and instrumentation specialists are essential to their development, it is necessary for the man himself whose problem is to be solved to grasp the methods which are to be used, or he will never appreciate their limitations or possibilities. It is therefore the task of those of us directly concerned with measurement work to encourage our fellow-engineers not only to make use of our new ideas, but themselves to undertake the practical solution of measurement problems which may arise in the course of their work.

Mr. Edmundson is with the British Thomson-Houston Co., Ltd.

Abstract No. 1851
Mar. 1955

SOUTH-WESTERN SUB-CENTRE: CHAIRMAN'S ADDRESS

By H. C. O. STANBURY, B.Sc.(Eng.), Member.

"TELECOMMUNICATIONS—A REVIEW OF SOME ASPECTS"

(ABSTRACT *of Address delivered at* PLYMOUTH, *14th October*, 1954.)

Local Aspects

The problem of developing the telephone system in an area which is largely rural lies in the extension of the local line plant rather than in the enlargement of the telephone exchanges. The Plymouth Telephone Area covers 1 850 square miles with 45 000 subscribers spread over 137 exchanges. Of these subscribers 17 300 are in seven urban territories, leaving 27 700 spread over 127 small exchanges in the rural areas, an average of only 220 subscribers per exchange. These exchanges have between them 2 500 waiting applicants, an average of 20 per exchange. The rural areas concerned were not telephonically very productive before the war, but since then, owing to changed relative circumstances and costs, there has been a considerable demand, particularly from remote farmsteads.

Post-war planning technique for telephone cable development includes the use of cabinets and pillars between the exchange and distribution points, the cable system being so designed that the smallest cables (15 pairs) between pillar and distribution points are laid to last for 20 years, because with cables of that size provision in instalments is uneconomic. The larger cables between cabinet and pillars are in general provided for 10 or 11 years, and those between exchange and cabinets for 5 or 6 years, with one or more additional instalments of cable scheduled for later in the 20-year period. Distribution-point cables are provided on the basis of a 70% fill, i.e. 10 pairs for every 7 subscribers expected, whereas pillar-and-cabinet cables are provided on the basis of a 90% fill. This is because the forecast demand for such a small unit as a distribution-point area must be more susceptible to error than is the case for large units of territory. Forecasts for exchange areas as a whole do prove in fact remarkably accurate. It follows that at each cabinet there will be terminated more pairs feeding out to the pillars than there are going back to the exchange, with a similar condition existing at each pillar where the aggregated distribution-point pairs will exceed the pairs going back to the cabinet. In both cabinet and pillar are facilities for connecting any pair on the exchange side to any pair on the subscribers' side, and the flexibility provided enables departure from the forecast in parts of the territory to be met without having to open joints and divert wires.

This technique is fully applicable in urban areas, but in sparsely populated rural areas of the size mentioned earlier, there is no place for a cabinet. The number of cable pairs terminated at the exchange itself is in fact no more than could be terminated on a cabinet. In many cases there is not even sufficient demand in any one direction to justify the provision of a pillar. Thus the small cable from the furthermost distribution point in any one direction is run right back to exchange, growing in size as other distribution-point cables are joined to it *en route*, all provided on the basis of 20 years and a 70% fill. It is thus clear that in terms of the number of subscribers likely to be connected in five years the scattered rural-exchange areas are more expensive initially than are the urban areas. But in the former, if the forecast is good, the plant once laid will last for 20 years, whereas in the latter additional instalments of cable will be required nearer the centre of the area during that period. Under limitations of capital expenditure, the most economic thing to do both from the point of view of early return for capital expended and of doing the greatest good for the greatest number is to spend most of the capital in urban areas. In a public service this has to be tempered by consideration for the position of people in rural territories, but a drive to clear up such areas in a short time would at its completion result in a redundancy of staff and a much reduced demand for capital, two factors which could contribute towards an undesirable deflationary effect. Thus there is still a waiting list, but a lot has been done. In the Plymouth Area subscribers are now being connected at twice the pre-war rate and the waiting list has fallen from 6 500 in 1949 to 4 500 in 1954. There are 45 000 subscribers compared with 25 000 in 1945, and trunk calls per month have increased from 223 000 in 1945 to 464 000 in 1954.

National figures show a similar trend, and the number of telephones per 100 population is 11·8, 70% greater than it was in 1939. The United States also have increased their penetration by 70% since 1942, though the degree of penetration is admittedly just over two and a half times our own.

Current Technical Developments

The transatlantic telephone cable to be laid as a joint British-American project is to make use of the American-type repeaters with flexible housings between Newfoundland and Oban. This type of housing enables the repeaters to pass round the cable drums, etc., with the cable, which can thus be laid without stopping the ship. But the smaller-diameter housing means that two cables must be used, one for each direction of transmission. In the Newfoundland–Nova Scotia section, for which the British Post Office will be responsible, the British type of repeater, with a rigid housing, will be used. This amplifies in both directions, with directional filters and equalizers on either side of the common amplifier, thus making possible the use of a single cable for transmission in both directions. All the cable laying will be carried out by H.M.T.S. *Monarch*, the largest cable ship in the world, with a gross tonnage of 8 056 and a loaded displacement of 14 000 tons. She was launched in August, 1945.

A major component in the development of trunk mechanization and the cordless switchboard is the motor-driven uniselector. This has 200 outlets and is very fast in operation. The motor drive is cut on seizing the required outlet by the operation of a high-speed relay which releases the latch magnet, thus dropping a latch into the gear wheel fixed on the wiper spindle.

A facsimile system of transmitting telegrams to small offices where the expense of a teleprinter and specially trained staff is not justified is under trial.

Mr. Stanbury is with the General Post Office.

DISCUSSION ON
"A SURVEY OF ELECTRICAL CERAMICS"*

SOUTH-WEST SCOTLAND SUB-CENTRE, AT GLASGOW, 4TH NOVEMBER, 1953

Mr. D. S. Gordon: In their mention of the ferromagnetic spinels the authors do not give values of permeability and resistivity for typical modern materials. It appears that the characteristics of these spinels represent a significant improvement on those of other core materials at frequencies from a few tens of kilocycles per second upwards. Some typical characteristics would therefore be of value.

No mention has been made of the piezo-electric properties of barium titanate and similar ceramics. Some very optimistic claims have been made for the sensitivity and general performance of electro-mechanical transducers employing these materials. On the other hand, more modest and apparently contradictory reports come from other sources. Could the authors give some authoritative figures for transducer elements made from available ceramics, so that their possible practical advantages might be assessed?

Mr. G. Robinson: I notice that, in general, high-permittivity ceramic materials of the barium-titanate type are formed from ions which are di- and tetravalent and which have a perovskite crystal structure. Have any high-permittivity ceramics been made from ions which are mono-, tri- or pentavalent, and have materials of a different crystal structure been found suitable?

In order to increase the capacitance of dielectrics, the form of stacked discs or plates is preferable to the tube form of ceramic, although it is more difficult to manufacture. I believe that the mineral clay *bentonite* (which can be prepared in the form of thin sheets) was proposed a few years ago, initially as a substitute for mica. It suffered from various disadvantages at that time. Has any recent work been carried out to improve its properties in relation to its use as a dielectric?

Mr. R. A. F. Bouette: In the manufacture of ceramic capacitors using rutile tubes, a problem is encountered sufficiently frequently to be troublesome. A typical capacitor would withstand successfully one, or several, tests at a voltage of up to 2kV—this being well under the ultimate breakdown voltage for the ceramic material used. When this capacitor is then operated at its normal working voltage of, say, 100–200 volts d.c., it would function normally for a time, and then, after probably several months of operation, it would break through the ceramic dielectric.

Could this be due to the initial test voltage causing a strain in the dielectric which ultimately leads to breakdown, or does the normal operation at working voltage, imposing a constant stress on the dielectric, gradually lead to an electrolytic change in the ceramic, which would again lead to breakdown?

Mr. T. Hall: A phenomenon which is noticed in the application of ceramics to r.f. circuits is "scintillation." The effect of this may readily be observed, but the literature available on the subject refers to it as random changes and fluctuations in the capacitor, and mentions rather vaguely the probable causes. Could the authors give any information on the subject?

Messrs. W. G. Robinson and **E. C. Bloor** (*in reply*): Mr. Gordon asks for further information on the ferromagnetic spinels. There are many materials now available with characteristics suitable for various purposes. Core materials for use at audio and high frequencies have initial permeabilities of 20–1 000 and maximum permeabilities of 100–4 000. Volume resistivities are of the order of 10^4–10^7 ohm-cm. The characteristics are frequency dependent, and the permeability falls to 2–20 at 10^8c/s. Other ferrites have been developed having high residual magnetism, for use as permanent magnets (retentivity up to 2 700 gauss), while others have rectangular hysteresis loops and are used as memory devices in computers.

The piezo-electric properties of barium titanate are dealt with in Reference 67 of the Bibliography, which refers to many original papers on this subject. In practice, the piezo-electric properties of barium titanate are profoundly affected by the presence of small amounts of other materials, and this probably accounts for the variable and contradictory reports to which Mr. Gordon refers. Barium titanate with the addition of a small percentage of lead titanate has given good results, and coupling coefficients of 0·1–0·5 are obtained.

In reply to Mr. G. Robinson, there are a number of materials with crystal structures similar to that of barium titanate which exhibit high permittivity, e.g. the columbates and tantalates of sodium and potassium which involve valencies other than 2 and 4.

Some of the ferrites which have the inverse spinel crystal structure also have high permittivities but are unsuitable for use as dielectrics on account of their low resistivity.

There has been some progress in the manufacture of ceramics in thin sheet form, the materials used being barium titanates with permittivities of 1 000–3 000. Bentonite is not now considered for this purpose on account of its low permittivity (about 6).

Mr. Bouette refers to the electrical breakdown of rutile materials after prolonged periods at comparatively low electrical stress. There are a number of possible explanations of this. The presence of small amounts of semi-conducting titanates or the reduction of titanium dioxide to the semi-conducting form Ti_2O_3 could lead to deterioration. Reduction of rutile can occur at temperatures as low as 600°C and may take place in the silvering process. This subject has recently been dealt with by W. A. Weyl and N. A. Terhune,† and it does not seem likely that the high test-voltage has any influence.

The phenomenon of scintillation as observed in ceramic capacitors appears to be due to the rather indefinite edge of the deposited silver electrodes. Further information on this subject may be found in E.R.A. Report Ref. L/T181.

* ROBINSON, W. G., and BLOOR, E. C., Paper No. 1452 M, March, 1953 (see 100, Part IIA, p. 247).

† WEYL, W. A., and TERHUNE, N. A.: *Ceramic Age*, 1953, 62, p. 23.

621.317.761.081.1

Paper No. 1765 M
Oct. 1954

THE STANDARD FREQUENCY MONITOR AT THE NATIONAL PHYSICAL LABORATORY

By J. McA. STEELE, B.Sc.(Eng.)

The paper was first received 17*th June, and in revised form* 13*th September*, 1954. *It was published in October*, 1954, *and was read before a joint meeting of the* RADIO *and* MEASUREMENTS SECTIONS 10*th November*, 1954, *and the* NORTH-EASTERN RADIO AND MEASUREMENTS GROUP 7*th March*, 1955.)

SUMMARY

Equipment has been constructed at the National Physical Laboratory for the automatic measurement and recording of the standard frequencies transmitted by MSF, Rugby, on 60 kc/s, 2·5, 5 and 10 Mc/s, and also the carrier frequency of Droitwich, 200 kc/s. For the latter and MSF 60 kc/s, the methods of measurement are an extension of those already well established for the intercomparison of the N.P.L. frequency standards. The standard deviation of measurements lasting a few seconds is 2–3 parts in 10^9 for MSF 60 kc/s, and 3–4 parts in 10^9 for Droitwich.

Measurements of the phase of the 1 c/s pulse modulation on MSF 60 kc/s can be made with very high precision. In daily comparisons, the scatter of a group of 60 readings does not usually exceed 30 microsec.

By a process involving frequency changing by continuous phase shifting, the received frequencies of MSF 2·5 and 5 Mc/s are recorded on a frequency meter of range ± 1 c/s, the discrimination being 2 parts and 1 part in 10^8 at these frequencies, respectively. The records obtained show well-defined diurnal variations on 2·5 Mc/s, particularly at sunrise, where the deviations may amount to several parts in 10^7. The received frequency of 5 Mc/s is subject to continuous variations over a range of about 1 part in 10^7 during daylight hours; in darkness larger changes are recorded.

(1) INTRODUCTION

A service of standard frequency transmissions has been provided from the Post Office Station, MSF, Rugby, since February, 1950. These transmissions are sponsored by the National Physical Laboratory, whose responsibility it is to ensure that the frequencies radiated do not depart from nominal by more than the stated tolerance of ± 2 parts in 10^8 when measured in terms of the frequency standard maintained in the Electricity Division. In addition to certifying the values of transmitted frequency, it is important also to investigate the variations induced by ionospheric transmission in the received frequencies of the several high-frequency transmissions in use and thus to evaluate the relative usefulness of these transmissions. Once the magnitude and incidence of these frequency changes has been determined, some guidance can be given to users as to the stability to be expected in the carrier and modulation frequencies and the periods of the day when the signals may be used with the highest accuracy.

It was to ensure the satisfactory discharge of these monitoring functions that the standard frequency monitor was constructed. In its design, emphasis has been laid on automatic, unattended, operation so that the necessary measurements may be carried out without undue demands on the time of the available staff. The description of the monitor falls naturally into three main divisions. The first is devoted to the measurement of the transmissions of MSF 60 kc/s and Droitwich 200 kc/s. The MSF transmission, which is made for one hour each day forms a link between the local frequency standard at Rugby and the frequency standard of the Laboratory, for the ground wave largely predominates in the received signal, and measurements of the received frequency may therefore be taken to apply also to the transmitted frequency. The carrier frequency of the transmission of the British Broadcasting Corporation on 200 kc/s is controlled within about ± 5 parts in 10^8 of nominal and is known to be widely used in this country as a secondary frequency standard. Its frequency is measured at the Laboratory continuously and automatically during all the hours of transmission.

Both MSF 60 kc/s and Droitwich are strongly received, and since their frequencies show little deviation from their nominal values, it has been found practicable to measure them by an extension of the methods used for the continuous and automatic intercomparison of the 100 kc/s oscillators in the frequency standard.

The second part of the paper is concerned with the methods for the precise measurement of the phase of the 1 c/s pulse modulation carried by all MSF transmissions for five minutes of each quarter-hour transmission. The highest accuracy in comparing the phase of the received pulses with pulses generated in the frequency standard is realized by use of the 60 kc/s transmission, and measurements have been largely confined to this frequency. With the advent of continuous high-frequency transmissions in May, 1953, the stability of pulse reception on these frequencies has also been investigated.

Finally, the measurement of the frequencies of the MSF transmissions on 2·5 and 5 Mc/s is described. Owing to the wide variations in received frequency and the fluctuating character of the signals, as compared with the low-frequency transmission, a fresh approach to the problem of frequency measurement was necessary. The method finally adopted has proved very satisfactory in practice and may be readily applied to other standard frequencies in the range 2·5–25 Mc/s. This part of the monitor has been in continuous operation since May, 1953. Examples of the results obtained are given and some of the more interesting effects are discussed.

(2) LOW FREQUENCY MEASUREMENT

(2.1) Manual Frequency Comparison at 100 kc/s

Before measurement, MSF 60 kc/s and Droitwich 200 kc/s are converted to 100 kc/s by mixing with suitable local oscillator frequencies in their respective receivers. Thereafter, both signals are treated for comparison purposes precisely as if they were generated locally and they take their place in the system of measurement at 100 kc/s along with the oscillators of the frequency standard. The following account of the method of frequency comparison applies equally, therefore, to received signal and local standard.

The relative frequency of any two 100 kc/s oscillators is found by beating them together and measuring the duration of one complete beat period. Since the measurement takes place in the elapsed time of one beat, what is involved essentially is a determination of the relative phase of the oscillations. This is accomplished in the phase comparator due to Law.[1] In this instrument the two frequencies are applied to a form of phase discriminator whose output falls to zero twice in each beat

The paper is an official communication from the National Physical Laboratory.

period at relative phase angles of about 90° and 270°. The beat period is defined by short pulses produced at successive passages through zero in the same direction. Law estimates that the error of the phase comparison is not greater than $\pm 0{\cdot}004°$, or about 1 part in 10^5 of the total beat period. This precision is at least ten times greater than is required in routine frequency comparisons on the monitor.

By means of circuits internal to the phase comparator, pulses denoting the beginning and end of the beat period are routed, respectively, to the "start" and "stop" terminals of an electronic counter of commercial design. This has five stages of decimal indication, and the time interval between the two pulses may be read from dials which show the total number of oscillations of the counter-driving frequency in the duration of the beat period. When the counter is supplied with a counting frequency which is a power of ten, the beat period is given direct in seconds and decimal parts of a second. The accuracy of any one measurement of time interval is limited to ± 1 unit in the last decade, or to 1 000, 100 or 10 microsec when the counting frequency is 10^3, 10^4 or 10^5 c/s, respectively.

For routine manual frequency comparisons two oscillators are connected to the phase comparator and a counting frequency is chosen which will yield the desired accuracy of measurement, depending upon the length of the beat period and the relative short-term stability of the oscillators. Thus, for a comparison between two of the principal oscillators in the frequency standard, the beat period is about 70 sec, and with a relative stability of about ± 1 part in 10^{10} a counting frequency of 1 kc/s is more than adequate. Shorter beats lasting about 2·5 sec occur frequently in other comparisons, with the same stability, and for these the counting frequency is increased to 10 kc/s. For these shorter beats ten consecutive measurements of alternate beat periods are taken. The resetting of the counter to zero after each measurement is accomplished automatically by a relay in the phase comparator. If so desired, the reset may be omitted and the readings allowed to accumulate on the register of the counter, until, with the tenth count, the mean beat period is obtained without further calculation.

(2.2) Automatic Measurement

To extend the method of comparison described above to automatic operation requires the provision of a recording counter and some system of switching which will present to the phase comparator pairs of oscillators for comparison according to some prearranged schedule. It is necessary, also, to supply the counter with the correct counting frequency for each comparison.

For automatic measurement it has been found possible to simplify the design of the counter considerably in virtue of the fact that all the sources of 100 kc/s, both local and received, are exceedingly stable and change their relative frequencies only very slowly. In most comparisons the change in beat period in the course of a week will not usually exceed 0·1% and may only amount to 5% in an interval of two months. Thus, once the total beat period has been measured, its full determination need not be repeated at every subsequent measurement. It is sufficient if the counter for use in the automatic monitor is able to accommodate only the changes in beat period within some suitable range which need be only a fraction of the full range of the decimal counter used in manual comparisons. In practice, the maximum register of the automatic counter is set at 0·25% of that of the manual instrument. A further point of difference in the counters is that the binary, as opposed to the decimal, scale is chosen for automatic recording. This change leads to rather more convenient and reliable operation of the counter.

In the course of a measurement of beat period, the recording counter will normally re-cycle many times, and as there are eight binary stages the number registered on receipt of the stop pulse is the remainder following the repeated subtraction of 2^8, or 256, from the total beat period. This number which lies, of course, between 0 and 255 is recorded not in its original digital form but in analogue form by producing a direct voltage which is proportional to it and using a recording millivoltmeter as the final indicator. The scale of this instrument is therefore a time scale with a maximum value of 255 periods of the counting frequency. The arrangement is shown in outline in Fig. 1.

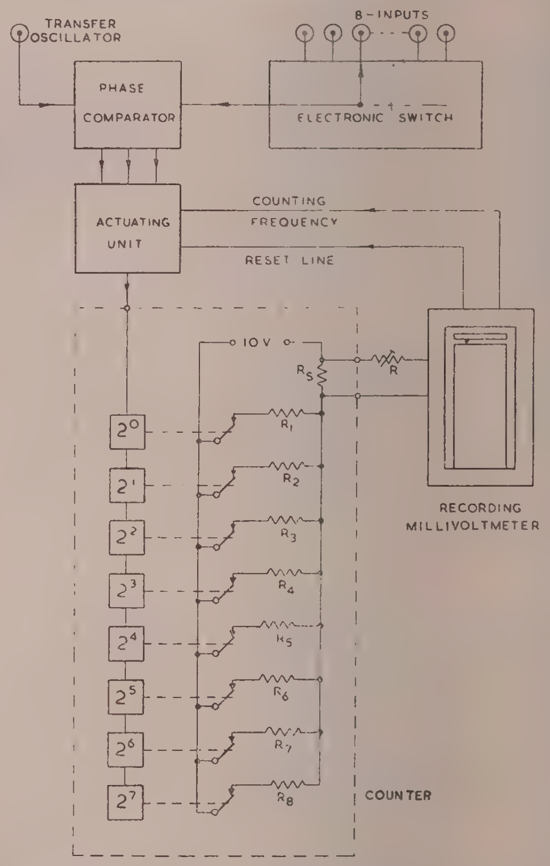


Fig. 1.—System of automatic frequency comparison at 100 kc/s.

The analogue of the counter readings is established by associating with each binary stage a pair of relay contacts. These are normally closed, and in this position the current flowing in each branch of the ladder network is determined by the series resistances R_1–R_8. The corresponding conductances are chosen to be in geometric progression with index of 2, i.e. $G_8 = 2G_7 = 2^2G_6 = \ldots = 2^7G_1$, and thus a scale of current is obtained in the eight branches of the network parallel to the scale of the counter. Each pair of contacts is caused to open when the stage associated with it has registered, i.e. for the nth stage when the count has reached the value 2^{n-1}, and so the change in current is proportional to the value of the corresponding counter stage.

The ladder network is supplied from a 10 volt stabilized source through a small series resistance R_S which is much smaller than R_8 the smallest of the series resistances. It acts, therefore, as a summing resistance for all the individual branch currents and the voltage across it is thus a measure, to some scale, of the total count registered by the counter. This voltage is applied in

series with the scale-adjusting resistance R to the recording millivoltmeter.

(2.3) Accuracy of Automatic Recorder

The scale of the recording millivoltmeter—or recorder, as it will be called—is a time scale, but it is more convenient in practice to use it direct as a scale of frequency difference. The frequency discrimination of the instrument is most easily seen by writing in the form:

$$\frac{\delta f}{f} = -\frac{\delta t}{t^2 f}, \text{ where } \frac{\delta t}{t} \ll 1 \quad . \quad . \quad . \quad . \quad (1)$$

δf, δt are the corresponding uncertainties in frequency and time, f is the nominal frequency of comparison and t is the time, in seconds, for one beat.

The left-hand side of eqn. (1) may be computed for given values of f, t and δt. The last is taken as one division of the chart of the recorder, and as this is ruled into 50 divisions, δt corresponds to 255/50, or $5 \cdot 1$, periods of the counting frequency. The frequency of comparison, f, is 100kc/s, and in the automatic recording of the N.P.L. frequency standards and of received transmissions, the range of t is about 2–5sec. Taking the above values of the parameters, $\delta f/f$, which may be called the scale factor, or change in frequency per chart division, is shown plotted in Fig. 2 for values of t of 1–7sec. In the majority

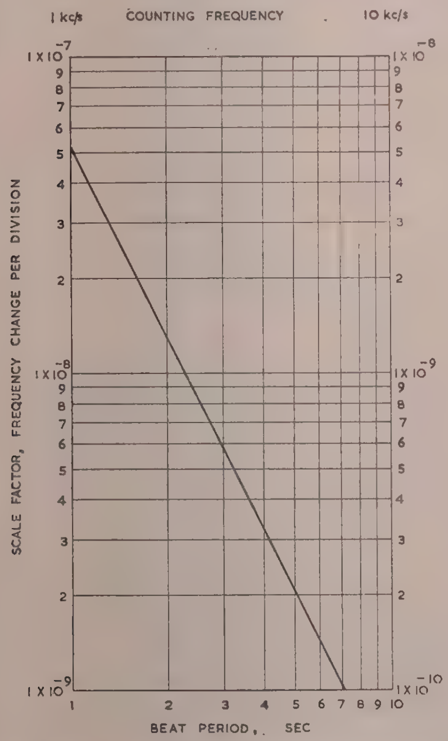


Fig. 2.—Recorder scale factor for counting frequencies of 1kc/s and 10kc/s.

of frequency comparisons either 1kc/s or 10kc/s is a suitable counting frequency and two scales are provided, one appropriate to 1kc/s when $\delta t = 5 \cdot 1$ millisec and the other to 10kc/s when $\delta t = 0 \cdot 51$ millisec. Other counting frequencies may be used, and there is no longer any advantage in restricting them to decimal powers as was convenient in manual comparisons, since the binary scale of counting is used; moreover, a scale factor intervenes in interpreting the readings of the chart record.

From Fig. 2 it will be seen that the scale factor varies over a wide range: the minimum value is about 5 parts in 10^8 for a beat of one second, counting at 1kc/s, rising to a maximum of 1 part in 10^{10} for a beat of 7sec, counting at 10kc/s. When the highest discrimination is required the counting frequency is increased to 50kc/s, yielding a scale factor of $1 \cdot 6$ parts in 10^{10} for a beat of $2 \cdot 5$sec, which is of common occurence. If the condition of eqn. (1) is satisfied, i.e. if $\delta t/t \ll 1$, the chart of the recorder is a linear scale of frequency difference and the scale factor applies unchanged across its width. It is not always possible to satisfy this condition, however, and then the scale factor appropriate to the mean position of the trace on the chart must be used. Thus if $\delta t/t = 1/10$, the scale factor varies by about 20% in passing from one side of the chart to the other. This non-linearity of the scale for some comparisons is not, in practice, a serious limitation, for the changes in beat period take place slowly and the total beat period is checked at intervals in routine manual comparisons. From these measurements the value of scale factor is found by means of Fig. 2 and the factor in use is adjusted, if necessary.

When applying Fig. 2 to the determination of scale factors for the received signals on 60kc/s and 200kc/s, the ordinates as given require adjustment. It will be clear later from the description of the receivers for these two frequencies that the error in cycles per second at 100kc/s is the same at either 60kc/s or 200kc/s. Therefore, for MSF, the ordinates of Fig. 2 should be divided by $0 \cdot 6$, and for Droitwich, by 2, to give the correct scale factor in each case, for changes in the received frequencies.

(2.4) The Recorder

In addition to recording the length of the beat period as registered by the recording counter at the end of each beat, the recorder performs a number of other functions. It is a multichannel instrument in which the movement is connected in turn to ten separate circuits by means of cam-operated, single-pole mercury switches, each of which remains closed for 18sec, the complete scan therefore occupying 3min. The position which the movement takes up in each 18sec period is recorded at the end of that time, when the pointer is forced into contact with the paper chart. The various channels are distinguished by interposing a multi-coloured ribbon between paper and pointer. The chart-speed is 15mm/h.

For the purposes of the monitor the recorder movement is separated from the mercury switches and is connected permanently across the summing resistor R_S of Fig. 1. All but one of the ten switches are used to apply the various counting frequencies to the recording counter. The time scale of the whole automatic measuring system is determined by the speed of rotation of the camshaft driving the switches, 18sec being available for each frequency measurement and the recording of the result. It is necessary in some comparisons to parallel two channels and thus double the time in which measurement may take place. Operated also by the camshaft are other switches performing necessary resetting operations every 18 or 36sec: the mercury switch not used for counting frequencies provides an overall reset to the measuring system at the end of every cycle of 3min.

In Fig. 1 it will be seen that the counting frequencies from the recorder pass to what is called an "actuating unit." This is an adaptation of a component part of the counter used for manual comparisons. It converts the sinusoidal waveforms supplied to it into square waveforms suitable for driving the counter. By means of gating circuits, controlled by the start and stop pulses

from the phase comparator, the counting frequency is applied to the counter for the duration of the beat period.

(2.5) The Electronic Switch

To complete the basic equipment for automatic measurement it is necessary to provide for the switching of a number of oscillators in turn to the phase comparator. In the system of measurement adopted in the monitor, one of the standard oscillators functions as a transfer standard and is connected permanently to one input of the phase comparator. Associated with the other input is what may be termed a single-pole multiway switch.

The principal requirement to be met in the design of the switch is that the leakage of the signal between channels shall not reduce appreciably the accuracy of comparison that can be obtained in the absence of the switch. This requirement is rendered less stringent by the fact that all the sources of 100kc/s normally connected to the switch occupy a very small range in frequency, not greater than 15 parts in 10^8. Considering crosstalk simply as frequency modulation of one oscillator by another, an attenuation between channels of 60dB is sufficient to reduce the maximum intermodulation effect to 15 parts in 10^{11}. This figure is of the same order as the short-period stability in any of the comparisons conducted on the monitor.

The required level of attenuation is realized in the switch by purely electronic means, using as the isolating element in each channel a pentode Type CV329 or CV2209. This valve has a short suppressor base of about 10 volts with normal potentials on the other electrodes. Eight valves are used in the switch with their anodes connected to a common load resistor. When operated at a screen potential of 20 volts, the residual signal at the output of the switch due to the effect of leakage from the sources connected to isolated channels is some 60 to 70dB below the normal signal level.

In the "on" and "off" positions of each pentode, potentials of +20 and −20 volts, respectively, relative to the cathode potential, are applied to the suppressor grid. These switching potentials are provided by the crystal diode matrix shown in Fig. 3. The

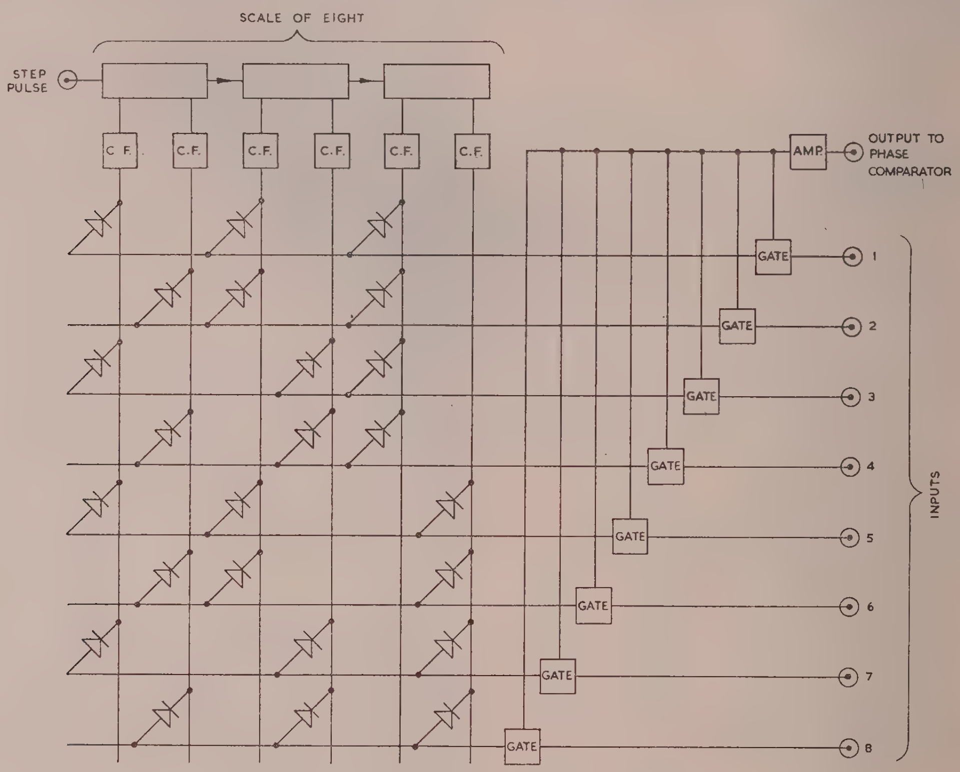


Fig. 3.—Schematic of 8-channel electronic switch.

purpose of the matrix is to relate the eight possible states of the scale-of-eight circuit to the eight possible positions of the switch. A fraction of the voltage at the anodes of each of the six valves in the scale is transferred by means of cathode followers to the six vertical busbars. The eight horizontal busbars terminate at one end on the suppressor grids of the gate valves, while at the other end each is connected by diodes to three of the vertical bars. For any configuration of the scale-of-eight, three of the vertical busbars are necessarily at a higher potential than the remainder. The connections of the diodes ensure that only one of the horizontal bars is attached to all three vertical bars at the higher potential and thus only one gate valve is turned on. All the other horizontal bars, being connected by one, or possibly two, diodes to vertical bars at the lower potential, remain tied to that value.

The successive positions of the switch are selected by the number of pulses received by the scale-of-eight. For example, when it has received five pulses, vertical bars 1, 3 and 6 are high in potential relative to 2, 4 and 5 and the fifth horizontal bar, being the only one connected to all three of the former, is therefore at +20 volts relative to the cathode of the gate valve. The input to channel 5 is thus the only one to appear across the common anode load and hence at the input of the phase comparator. In automatic operation the operating pulses are sup-

plied by the actuating unit, while for manual comparisons a telephone dial selector is used to produce a preset number of pulses for the selection of any desired channel.

(2.6) Organization of Frequency Comparison at 100 kc/s

The component parts of the measuring system having been described, it may be helpful to supplement Fig. 1 by following through a complete cycle of operation.

Assuming that an overall reset has been applied to the system by means of the switch in the recorder, the input on channel 1 of the switch is now connected to the phase comparator, and beats with the transfer oscillator. The required counting frequency is supplied to the actuating unit and the beat period between the oscillations is measured by the recording counter. Coincident with the receipt of the stop pulse from the comparator, the actuating unit pulses the electronic switch to channel 2. A fresh oscillator is counted to the phase comparator, but no measurement of beat period is possible until the end of the first 18 sec period, when the number held in the counter is registered on the chart of the recorder. Until this has been done and the counter has been reset, the actuating unit remains insensitive to the sequence of start and stop pulses from the phase comparator.

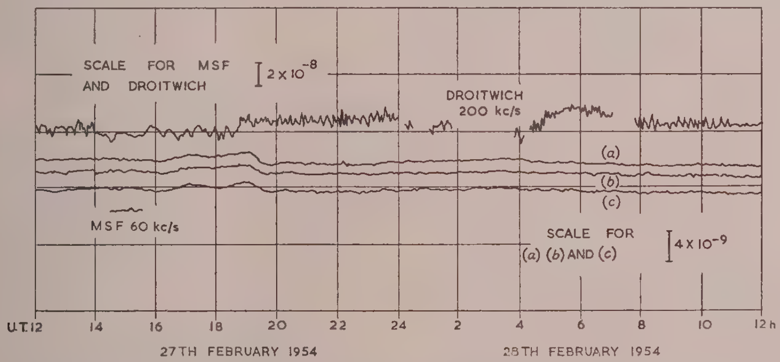


Fig. 4.—Record of 100 kc/s frequency recorder for a period of 24 hours, commencing 1200 U.T. on 27th February, 1954.
MSF 60 kc/s is recorded between 1430 and 1530 U.T. on the first day and Droitwich 200 kc/s during all its hours of transmission. Traces (*a*), (*b*) and (*c*) are intercomparisons of N.P.L. standards.

The first record having been made, the actuating unit is reset, the second frequency comparison proceeds, and the result is recorded at the end of a further period of 18 sec.

A series of comparisons is thus made and recorded until all the inputs to the switch have been measured relative to the transfer oscillator. On the first unoccupied channel of the switch the process comes to a halt, there being then only one oscillator connected to the phase comparator. No further comparison is possible until the overall reset is applied at the end of the 3 min cycle. Since a vacant channel interrupts the progression of measurements it is necessary to place the output of the Droitwich and MSF receivers on the later channels of the switch, after the oscillators in the frequency standard, so that the comparison of the latter may proceed at all times. Likewise, Droitwich precedes MSF in the switching sequence since its transmitting period embraces that of MSF.

(2.7) Results at 100 kc/s

A section of the chart of the frequency recorder covering a period of 24 hours commencing at midday is reproduced in Fig. 4. The records of the received frequencies of MSF and Droitwich will be fully discussed in a later Section. Of the remaining traces, those labelled (*a*), (*b*) and (*c*) are the records of the comparison of the three high-precision ring oscillators, which form the basis of the frequency standard of the Laboratory, with the transfer oscillator. The latter does not possess the high stability of the ring oscillators, but it is useful in the operation of the automatic equipment, since, being 4 parts in 10^6 from nominal, it gives a conveniently short beat with all the other oscillators and received signals. The transfer oscillator, as its name implies, is not used for absolute measurements. All measurements made in terms of it are transferred by means of the chart records to one of the ring oscillators for final determinations of frequency. The scale factor for the records (*a*), (*b*) and (*c*), for which the counting frequency is 10 kc/s, is about 8 parts in 10^{10} per division, and it is clear from the sympathetic changes in all three records that the source of variation lies largely, not in the rings, but in the common oscillator.

(2.8) The Reception of MSF 60 kc/s and Droitwich 200 kc/s

High field-strengths are provided by both MSF 60 kc/s and Droitwich, and the receiving equipment need not therefore be elaborate. The signals are received on a common aerial, a large inverted-L, 250 ft long and about 80 ft high. It is situated some 200 yd from the measuring station, to which it is connected by underground coaxial cable. A cathode follower is used as an impedance transformer between aerial and cable.

The received signals are converted to 100 kc/s in specially constructed, fixed-frequency receivers. Fig. 5 is a block schematic of the receiver for MSF. The receiver for Droitwich is similar

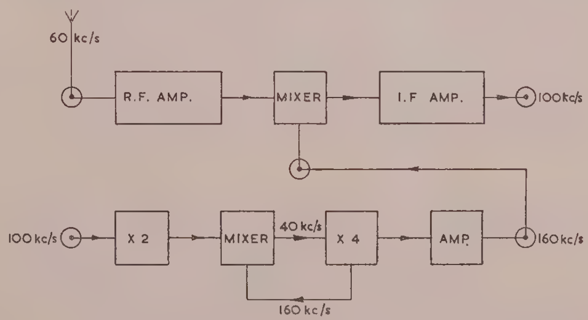


Fig. 5.—Schematic of receiver and local-oscillator generator for MSF 60 kc/s.

in general design. Following two stages of signal amplification, the received frequency is mixed with 160 kc/s in the case of MSF and 300 kc/s for Droitwich to yield a final output at 100 kc/s in both receivers. The local oscillator at 160 kc/s is obtained by regenerative multiplication as shown in Fig. 5: the 300 kc/s for Droitwich is produced by straightforward multiplication from

100kc/s. A narrow-band crystal filter may be inserted after the mixer to reduce the level of unwanted modulation or noise.

(2.9) The Measurement of MSF 60kc/s and Droitwich 200kc/s

Before discussing the accuracy of the measurements of the received frequencies of MSF and Droitwich, it is worth while recalling the method of measurement and emphasizing that the value obtained in the frequency comparison is a mean value averaged over the beat period. This point is well illustrated by an occurrence shortly after the start of MSF transmission in 1950. A user reported that he was unable, after removing all known sources of variation in his equipment, to adjust a local standard of 3Mc/s to zero beat with the 50th harmonic of MSF 60kc/s. The MSF carrier, it appeared, was subject to more or less regular phase variations of amplitude approximately 0·07° at a repetition frequency of about 1c/s. At this time, measurements were being conducted at the N.P.L. with a beat period of about 8sec, and consequently the maximum effect which this modulation could produce was to vary the phase by 0·14° in this period, corresponding to a change in mean frequency of roughly 2 parts in 10^9. This was only slightly more than the normal accuracy of measurement, and the modulation went undetected until means were devised for "looking" at the phase of the carrier over a very much shorter period, when the presence of the effect was confirmed. The Post Office identified the source of the trouble at the transmitter and steps were taken to remove it.

An example of the measurement of the received frequencies of MSF and Droitwich has already been seen in Fig. 4. MSF is recorded during its one-hour transmission, commencing at 1430 U.T. The beat period for this record is 4·0sec, and the scale factor 5 parts in 10^9 per division. The scatter of the points falls within the limit of about one division, and the mean frequency during the hour may be estimated with an error of less than 2 parts in 10^9. Similar results are obtained in manual measurements where the standard deviation of a group of ten readings is generally between 2 and 3 parts in 10^9.

Droitwich is recorded with a beat period of 2·2sec, the scale factor being the same as for MSF, namely 5 parts in 10^9 per division. The short-period scatter extends to some two divisions, or 1 part in 10^8, and during the 24h period the frequency varies over a range of about 5 divisions, or 2·5 parts in 10^8. In manual comparisons, the standard deviation of a series of ten measurements usually lies between 3 and 4 parts in 10^9.

There is some gain in accuracy in the use of longer-beat periods, and both transmissions are compared directly with one of the ring oscillators. The beat period for MSF is 170sec, and for Droitwich about 50sec. Three readings are taken and the maximum difference does not often exceed 1 part in 10^9.

(3) THE MEASUREMENT OF THE PHASE OF THE 1c/s MODULATION

(3.1) Method of Measurement

Pulse modulation at 1c/s is applied to all MSF transmissions in the second five minutes of each quarter-hour period. The pulses take the form—now widely adopted for standard frequency transmissions—of five complete sinusoidal oscillations at a frequency of 1kc/s, commencing with a positive half-cycle. All the pulse transmissions are received on a communication-type receiver. The choice of the relatively large pass band of ±750c/s, although it increases the effect of noise and interfering signals, ensures that the distortion of the pulse waveform is small; it also reduces the delay through the receiver.

As received, the form of the pulse is not well suited to precise measurement, and it is necessary to specify a reference point in the pulse length of 5millisec. The point chosen is that marked A in Fig. 6, which is a representation of the first three half-cycles of the pulse waveform. By means of the circuit described below, a pulse of duration a few microseconds is produced at this point. The pulse is applied to the start circuit of a decimal counter similar to those used for manual frequency comparisons. To the stop circuit of the counter are connected permanently pulses at a repetition frequency of 10c/s generated from one of the standard oscillators. The counter is supplied with 100kc/s counting frequency, and so the delay between received and standard pulses, which lies in the range 0–0·1sec, may be measured with a reading accuracy of 10microsec.

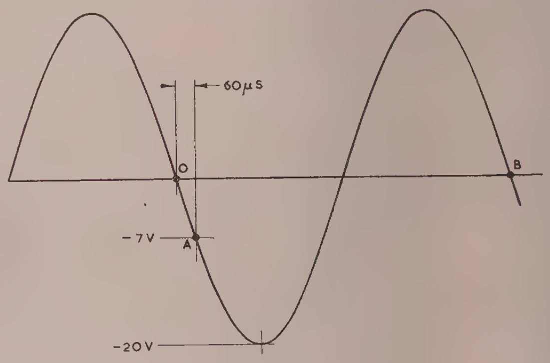


Fig. 6.—First three half-cycles of 1c/s pulse waveform.
The trigger point is marked A.

An amplitude comparator of the "multiar" type[2,3] generates the short pulses at the reference point: the circuit is shown in Fig. 7. The feedback path of the blocking oscillator is completed through diode V_1 in series with the secondary of the receiver output transformer, one side of which is returned to the slider of the voltage divider R. The latter provides a positive bias to the cathode of the diode: in the absence of a signal the diode is non-conducting, the feedback path is broken and the circuit is therefore quiescent.

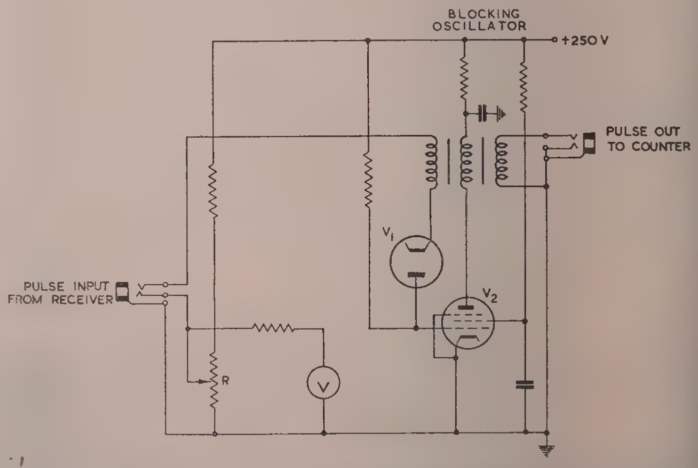


Fig. 7.—Circuit of "multiar" amplitude comparator used in pulse measurements.

On receipt of a pulse the cathode of V_1 is carried negative in the second half-cycle of the pulse waveform, and when the amplitude of the pulse is sufficiently great to cancel the bias potential the diode conducts, regeneration occurs and a sharp pulse is obtained from the output of the blocking oscillator. More than one pulse is produced, since the time-constants associated with the grid circuit of V_2 are very short, but those additional pulses are without effect on the counter, the start circuit of which is inactive following its operation by the first pulse. The cathode potential of V_1, and with it now the grid potential of V_2, since the diode is conducting, continues to fall with the signal potential until finally V_2 is cut off when the pulse amplitude is sufficiently large. Additional pulses are generated when the pulse potential passes through the grid base of V_2 in succeeding cycles, but, as before, these are irrelevant to the further triggering of the counter.

To ensure high precision in triggering the multiar, a large pulse amplitude of 40 volts peak-to-peak is used. The slope at the point A is thus 130 volts/millisec, or 1·3 volts per 10 microsec, i.e. 1·3 volts for a change of one place in the last decade of the counter. The largest source of circuit instability is presumed to lie in variations in the heater voltage of V_1. Any such effect must, however, be small, since later designs of the circuit using a crystal diode for V_1 have shown no significant difference in performance. Another source of error is variation in pulse amplitude. As indicated in Fig. 6 the delay between O and A is about 20° in phase or 60 microsec in time. If the counter reading is not to change by more than one unit in the last place, amplitude variations must not exceed about 20% of the peak value. This requirement is easily satisfied for reception of MSF 60 kc/s, but it is not always possible to ensure such stability in measurements on the high-frequency transmissions. There, however, variation in pulse amplitude is only one of several factors contributing to the overall accuracy of comparison.

(3.2) Operational Procedure

Prior to a measurement, the voltmeter V (Fig. 7) is set by means of R to read 20 volts, and the gain control of the receiver is varied until the received pulse just triggers the multiar at this value of bias. The voltmeter reading is then reduced to 7 volts, corresponding to the point A. This setting has been found to present a suitable compromise between the effect of variations in amplitude and the possibility of triggering by noise pulses or interfering transmissions. A certain number of spurious readings are still obtained at this bias, and should exceptionally noisy conditions prevail, the voltmeter may be set to a higher bias and correction made so that the readings will be comparable with those taken at the normal bias setting.

The use of 10 c/s pulses to stop the counter allows readings to be taken every second since there is sufficient time to note the reading and reset the counter to zero between successive received seconds pulses. A practised observer is able quickly to form a visual estimate of the mean reading, disregarding obviously false results. It is possible, also, to establish a rhythm in the operation, delaying the counter reset until just before the start pulse is received and thereby reducing the possibility of spurious readings. In determining the relative phases of pulses received from different sources, which may differ in phase by more than 100 millisec, any ambiguity of reading may be resolved either by listening to the pulses and estimating the time interval between them, or, more satisfactorily, by substituting 1 c/s pulses for the 10 c/s ones normally fed to the counter.

(3.3) Results on 60 kc/s

In daily measurements of the 1 c/s modulation on MSF 60 kc/s by the method described above, a high accuracy of comparison is consistently obtained. The measurement is confined to the last minute of the 5 min modulation period, and the total scatter of the readings does not usually exceed 30 microsec. On days which are exceptionally free from interference, the limits of discrimination of the counter are reached, the readings alternating between two places in the last decade. Thus, the error of measurement may with certainty be put at less than ±15 microsec in the great majority of comparisons. It is necessary to confine the readings to a period of about one minute, otherwise the rate of the standard supplying the 10 c/s pulses must be taken into account, the oscillator generally employed in pulse measurement having a rate relative to MSF of about 10 millisec/day or 6 microsec/min.

As a check on the internal accuracy of the comparison, two sets of readings are taken, 15 min apart. The relative rate of signal and standard per quarter-hour is 1/96 of the daily rate, and the difference between the two sets of readings should therefore be about 1% of the difference between readings on that day and the previous day. This is usually so to within 10%.

(3.4) Results on the High-Frequency Transmissions

The limitations of the pulse-measuring equipment described above in respect of fading signals may be largely overcome in practice by selecting the highest in a series of readings. Apart, however, from any deficiencies in the equipment, the stability of the pulse phase on 2·5 and 5 Mc/s cannot be expected to equal that obtained on 60 kc/s. Ionospheric transmission introduces variation in both signal path and velocity, while the length of the pulse does not permit signals arriving by different paths to be separated. No systematic daily measurements have been made on the high-frequency transmissions, but the results so far obtained indicate that a useful stability of about 0·1 millisec may be realized. A higher precision of comparison is possible, but it is doubtful whether it can be translated into a corresponding accuracy of measurement. Photographic records of the pulses received on 5 Mc/s show that considerable distortion of the pulse waveform may occur, and that it is therefore advisable to monitor the pulse waveform when carrying out precise comparisons.

(4) THE HIGH-FREQUENCY MSF TRANSMISSIONS

(4.1) General

Provision has been made for simultaneous MSF transmissions on three of the standard frequencies of 2·5, 5, 10, 15 and 20 Mc/s. So far, only the first three frequencies have been used, and of these, only 2·5 and 5 Mc/s are received regularly in Teddington at the present phase of the sunspot cycle. The operation of the high-frequency monitor has therefore been confined to these frequencies although receivers are available for all five frequencies and also for 25 Mc/s.

As mentioned earlier, the wide variations in the frequency and amplitude of the received high-frequency signals have necessitated a different method of recording from that adopted for 60 kc/s and 200 kc/s: it may be applied equally well to all the standard frequencies.

(4.2) Reception of the Standard Transmissions

As an example of the general design of the high-frequency receivers, that for 5 Mc/s is shown at the bottom of Fig. 8. Beyond the first stage of frequency changing, all six receivers are identical. A double frequency-change is employed to reduce the received frequency to 100 kc/s. Following two stages of amplification, the 5 Mc/s signal is mixed with 6·6 Mc/s to produce the first intermediate frequency of 1·6 Mc/s. This is amplified in

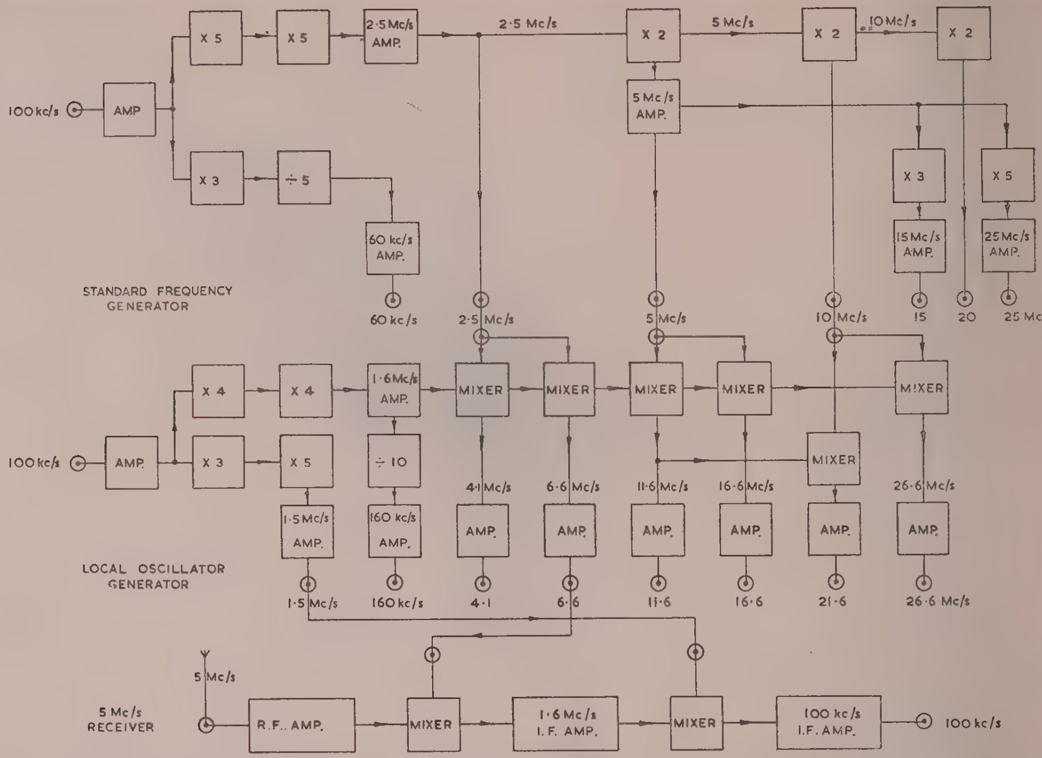


Fig. 8.—Standard and local-oscillator frequency generators for high-frequency receivers with schematic of receiver for 5 Mc/s.

four stages and then mixed with 1·5 Mc/s to give the second intermediate frequency of 100 kc/s. A crystal filter can be switched into circuit after the second mixer to reduce the bandwidth at the half-power level from the normal value of about $\pm 2\,000$ c/s to approximately ± 100 c/s. Full receiver output voltage is obtained with a signal of rather less than $1\,\mu$V. Delayed automatic gain-control is applied over the r.f. and first i.f. stages. In addition to the crystal-filter switch, two gain controls are provided on the front panel: one operates in series with the automatic gain-control and the other controls the level of the 100 kc/s output.

For reception of the standard frequency signals, a form of vertical rhombic aerial is used. It is mounted at a height of about 60 ft, roughly broadside to the direction of Rugby (45° W), and serves also for the reception of the transmissions of WWV from the same general direction (72° W). The aerial is situated in the open, some distance removed from sources of locally-generated noise and is connected to the receivers via an amplifier and underground cable.

(4.3) Standard and Local-Oscillator Frequency Generators

All the receiver local-oscillator frequencies are derived from one of the 100 kc/s oscillators in the frequency standard. Their generation takes place in two units, one of which, the "standard frequency generator," produces the local equivalents of the received frequencies in the range 2·5 to 25 Mc/s; the other, the "local oscillator generator," making use of some of these frequencies, supplies all the local oscillator frequencies—from 1·5 to 26·6 Mc/s. The arrangement is shown in schematic form in Fig. 8.

All the standard frequencies, it will be seen, are obtained by a process of multiplication from 100 kc/s: in the local oscillator generator both multiplication and mixing are used to obtain the final output frequencies. As an illustration, the 15 Mc/s receiver requires two local-oscillator frequencies—16·6 and 1·5 Mc/s. To produce the former the controlling 100 kc/s standard is multiplied by 4 and 4 to 1·6 Mc/s: this frequency is then mixed twice with 2·5 Mc/s to yield 6·6 Mc/s, whence, by successive mixing with 5 Mc/s, the desired first local-oscillator frequency is obtained. 1·5 Mc/s is common to all receivers and is generated by multiplication by 3 and 5 from 100 kc/s. A crystal filter is used to remove any residual modulation at 100 kc/s before supplying the receivers from eight low-impedance outputs.

The possibility of leakage from the standard frequency generator to the receivers is apparent and is guarded against by arranging that the signal frequency is not required in the production of the local-oscillator frequencies for any receiver except in the case of 2·5 Mc/s, where the frequency of the first local oscillator is the sum of signal and i.f. frequencies. A switch is incorporated in the standard frequency generator which allows the stages of multiplication beyond that required for the receiver in use to be disconnected.

(4.4) The Method of Frequency Recording

The signals received on 60 kc/s and 200 kc/s display good amplitude stability and high phase constancy, and the variations in frequency from day to day are nearly always less than about 1 part in 10^9 for MSF and 1 part in 10^8 for Droitwich. The received high-frequency transmissions, on the other hand, are generally subject to some measure of fading, sometimes very severe; the phase shows, frequently, rapid short-period variations, while the frequency regularly undergoes changes of several parts in 10^7. This essential variability in the received signals makes it difficult to perform frequency measurement by the relatively

simple method used for the low-frequency transmissions, and it has been necessary to develop equipment better adapted to the characteristics of the received signals.

In the h.f. recorder, beats are still formed between the output of the receiver and a standard source of 100 kc/s. They do not, however, take place around zero frequency, as in the 100 kc/s recorder, but in the region of 25 c/s. Furthermore, variations in the received frequency are not measured as variations in beat period but direct as frequency changes within the range 24–26 c/s. These limits are the extremes of the scale of the power-frequency recorder which is used as the final indicator of frequency difference between signal and standard. The calibration of the recorder is known to 0·05 c/s, and the received frequencies of 2·5 Mc/s and 5 Mc/s can therefore be measured to an accuracy of 2 parts and 1 part in 10^8, respectively.

The general system of measurement is shown in outline in Fig. 9. In the receiver the higher-frequency stages remain as

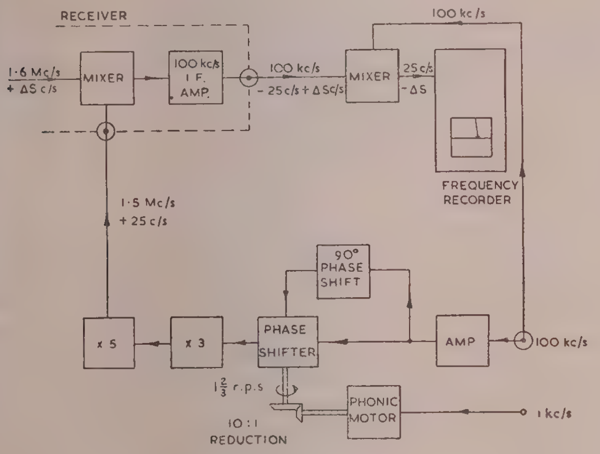


Fig. 9.—System of high-frequency recording.

already described and alteration appears only at the second local oscillator. It is necessary that the output of the receiver shall differ from 100 kc/s by 25 c/s, and this is achieved by increasing the frequency of the local oscillator to 1·5 Mc/s + 25 c/s. The increased frequency is still obtained by multiplication from the 100 kc/s standard by 3 and 5 as already described, but a continuously-rotating phase shifter is introduced in the input to the multiplier. It rotates synchronously at $1\frac{2}{3}$ c/s, and in passing through it the frequency of the standard is increased from 100 000 c/s to 100 001$\frac{2}{3}$ c/s. Thence, by multiplication, the final frequency of 1 500 025 c/s is obtained. By this device it is still possible to measure received frequencies direct in terms of that of one of the principal oscillators without recourse to specially off-set sources requiring frequent calibration.

The deviations ΔS of the signal frequency are transferred to the output of the receiver at a frequency of 99 975 c/s, and after mixing with standard 100 kc/s, they appear as variations about the centre frequency of the recorder.

(4.5) The Phase Shifter

The success of the method of measurement hinges on the satisfactory performance of the phase shifter. This instrument is required to advance the phase of the 100 kc/s input in as uniform a manner as possible at the rate of $10\pi/3$ rad/sec. It is important that departures from a linear relation between phase angle and angular rotation be kept small, for the effect of any phase modulation is increased by a factor of 15 in the subsequent multiplication.

By way of example, a sinusoidal variation from a linear characteristic of amplitude 0·6° and frequency $1\frac{2}{3}$ c/s will appear as equivalent frequency modulation of $\pm 0{\cdot}25$ c/s at 1·5 M/cs. Such a variation would embrace a total of 10 divisions of the chart of the frequency recorder.

The phase shifter used is of the electromagnetic type[4] in which a moving coil revolves in the crossed fields produced by two fixed coils displaced 90° in space phase and supplied with currents differing by 90° in time phase. All three coils are wound in slots in polystyrene formers with dust cores, the distribution of the turns in each phase of the stator winding being such as to produce a uniform field at the rotor. A circular field rotating at 100 000 r.p.s. is set up by the combined action of the field coils. When the rotor is stationary the e.m.f. induced in it has the same frequency as the supply: on rotating it against the direction of rotation of the field a higher frequency is obtained; conversely, a rotation in the same direction as the field produces a lower rotor frequency. In practice, owing to inevitable inaccuracies in construction, to small departures from equality in the magnitude of the field currents, and to small defects from 90° in their relative phase angle, the rotating field has an elliptical form and the output of the phase shifter is modulated in both amplitude and phase. The latter produces frequency modulation directly, while the former induces secondary frequency-modulation by varying the phase of the signal at the grid of the first multiplier. The magnitude of both forms of modulation is kept small by the method adopted for the control of the currents supplied to the field coils of the phase shifter. The circuit is shown in Fig. 10.

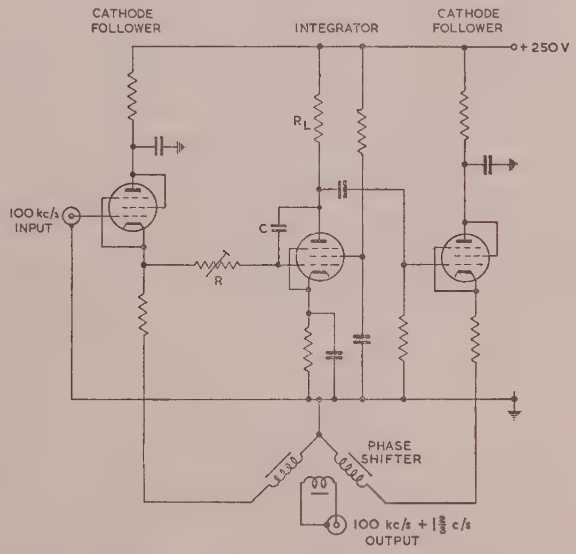


Fig. 10.—Phase-splitting circuit for 100 kc/s phase shifter.

The design follows that of a similar phase shifting network for a frequency of 80 kc/s[5]. One phase of the stator is supplied by a cathode-follower direct from the 100 kc/s source. The necessary difference in phase of 90° is introduced in the current supplied to the other phase by means of a form of Miller integrator in which the circuit constants are so adjusted that, for a frequency of 100 kc/s, the output voltage is equal to the input and leads it by 90°. To produce this result the following relations must hold simultaneously:

$$R_L = \frac{1}{\omega C}$$

and
$$R = R_L \frac{g_m R_L - 1}{g_m R_L + 1} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where g_m is the mutual conductance of the valve, $\omega = 2\pi \cdot 10^5$ rad/sec and R, R_L and C are as indicated in the Figure. If the first equation is satisfied to within 2%, the error in phase angle between output and input will not exceed a few minutes of arc. Equality of input and output voltages is obtained by adjustment of R, and, in practice, this is set to realize minimum frequency-modulation of the trace on the recorder. The residual modulation amounts to about $\pm\frac{1}{2}$ division at a repetition frequency of $1\frac{2}{3}$ c/s.

The synchronous drive to the phase shifter is provided by a phonic motor supplied with standard 1 kc/s. A worm gear reduces the speed of the motor from 1 000 r.p.m. to 100 r.p.m. at the shaft of the phase shifter. The gear ratio is not precisely 10 to 1 but about 2% less, so that an increment of 0·5 c/s is added to the local oscillator frequency. This compensates for the fact that the standard oscillator is about 15 parts in 10^8 below nominal and ensures that received signals are more symmetrically disposed about the mid-scale frequency of the recorder.

(4.6) The Frequency Recorder

The frequency recorder used is of the Shotter pattern.[6] In its principle of operation it resembles a differential induction wattmeter. What may be termed the current coils are tuned to different frequencies, one above and the other below the recorder frequency range. At some intermediate frequency the torques which the coils exert on the moving system in conjunction with the voltage coil, which is untuned, are arranged to balance. This frequency is the centre frequency of the recorder, and for a short range on either side, ± 1 c/s in the present instrument, a linear frequency scale is obtained by a suitable choice of the geometry of the moving system. The nominal supply voltage is 110 volts, and the burden of the recorder is 10 VA at 0·5 power factor. The necessary power is supplied by an audio amplifier.

One valuable feature of the recorder which commended it for use as the final indicator is that the indication is substantially independent of voltage for a variation of $\pm 10\%$, while for changes of $\pm 20\%$ the error is $\pm\frac{1}{2}$ division. There is thus no need to limit the output voltage of the receiver beyond the regulation provided by the operation of the automatic gain-control. The large reductions in signal amplitude which occur during fading do not produce appreciable error as the power input to the recorder falls below the level required to drive the movement. The pen therefore remains stationary at the last recorded value, resuming its indication when the signal strength is restored.

The inherent damping of the movement is increased by an adjustable oil dashpot. A suitable value of time-constant for general recording is found to be 5 sec. This permits the indication of even large, abrupt frequency changes to be substantially complete in the shortest time that can be readily distinguished on the record. The chart speed is 3 in/h, and the limit of discrimination is about 30 sec. The additional damping also reduces the extent of the residual frequency modulation of the trace to about half a division, i.e. 1 part in 10^8 at 2·5 Mc/s or 5 parts in 10^9 at 5 Mc/s. These figures represent the useful discrimination of the recorder.

One point in the calibration of the scale is checked periodically. For this purpose the receiver in use is supplied with the signal frequency obtained from the standard frequency generator. The difference between local and "received" frequencies would be zero but for the addition of 0·5 c/s to the local oscillator frequency in the reduction gear. The calibration therefore appears at $+0\cdot5$ c/s relative to the recorder centre frequency. One of these calibration lines extending over about 20 min will be seen in Fig. 12.

(4.7) Records at 2·5 and 5 Mc/s

In Figs. 11 and 12 an attempt has been made to convey the essential features of the frequency records for one day on 2·5 and 5 Mc/s, respectively. It has not been possible to reproduce the short-period fluctuations in received frequency, but the "ripples" in some parts of the records are intended to indicate the magnitude, if not the true periodicity, of these variations. It is to be noted, also, that the MSF transmissions are interrupted between 15 and 20 min past each hour. The movement of the recorder is entirely under the control of the signal, and in its absence the pen will normally trace a straight line from the last recorded frequency. At times, however, interfering stations may continue to supply a signal sufficient to deflect the movement and an apparent frequency will be indicated until MSF is restored.

In Fig. 11 the marked division into day- and night-time con-

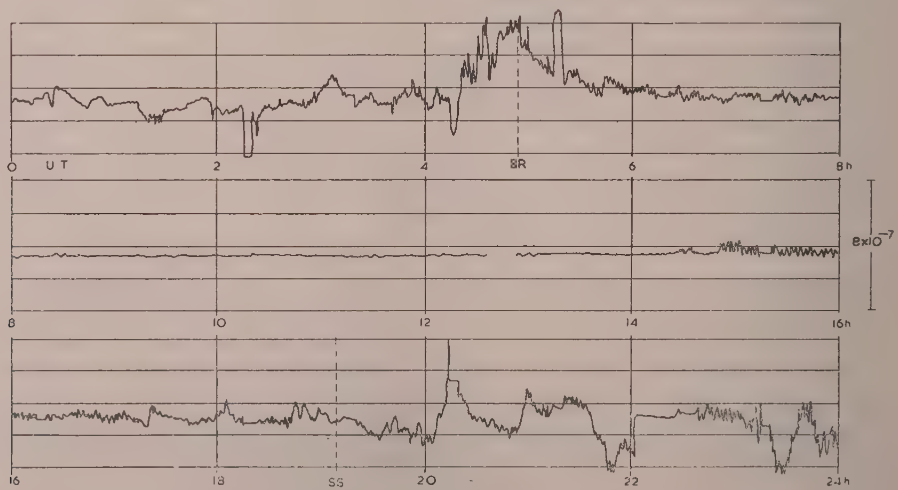


Fig. 11.—Received frequency variations on 2·5 Mc/s for three 8-hourly periods on 23rd August, 1953.

A deflection of the trace upwards corresponds to an increase in received frequency.

ions will be apparent. The received frequency is most stable the period between 0800 and 1430 U.T., when the variations nount to about 3 or 4 parts in 10^8, although there are sections the record in which the trace is, for a short time, sensibly a aight line, indicating that the stability of the received frequency better than 5 parts in 10^9. The pronounced increase in freency of some 4 or 5 parts in 10^7, reaching the upper limit of e chart in the sunrise period, is the result of the transfer of flection from the F- to the E-layer. After sunset, large negative equency changes are indicated, twice extending to the lower nit of the record. The reversion to the normal day-time ability for a period of about 40 min just after 2200 hours is due the sudden onset of a strong sporadic-E reflection, when the tensity of ionization in the E-layer is raised locally and temprarily above the usual value for the time and season.

The interpretation of the record for 5 Mc/s in Fig. 12 is com-

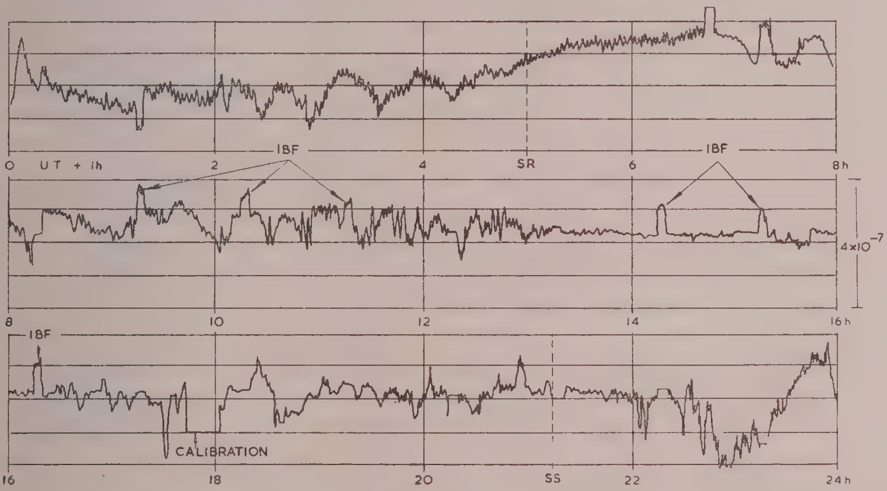


Fig. 12.—Received frequency variations on 5 Mc/s for three 8-hourly periods on 14th July, 1953.
A deflection of the trace upwards corresponds to an increase in received frequency.

plicated by the reception in the early-morning hours of station VWV of the National Bureau of Standards. WWV 5 Mc/s is trongly received, as is shown by the absence of the MSF 5 min ilent periods; the received frequency is high by about 1 part 1 10^7 at sunrise following the progressive establishment of ower levels of signal reflection. During daylight hours the uctuating record of MSF 5 Mc/s corresponds to the known isturbed nature of the F-layer. The recorded changes in eceived frequency occupy a range of about 1 part in 10^7, and he longer fluctuations have a quasi-period of between 10 and 0 min. As with 2·5 Mc/s, strong sporadic-E reflections display ood frequency stability. A striking example occurs between 300 and 1500 hours, when the received frequency varies by only or 2 parts in 10^8.

A third source of standard frequencies in Europe on 5 Mc/s re the transmissions of station IBF, Turin, between 0800 to 100 and 1300 to 1600 U.T. on each Tuesday. The frequency of his station is shown recorded in Fig. 12 during the six MSF silent eriods which fall within its transmitting times.

(5) ACKNOWLEDGMENTS

The method of frequency comparison at 100 kc/s follows enerally the practice of the Post Office Radio Experimental and Development Laboratory, which also supplied the three Essenring crystals that form the basis of the frequency standard. The British Electricity Authority rendered valuable assistance in the early stages of the development of the high-frequency monitor by the loan of a frequency recorder. All the high-frequency records reproduced in the paper have been obtained with a frequency recorder kindly lent by Messrs. Elliott Brothers Research Laboratories, to whom grateful thanks are extended.

The Metrology Division of the N.P.L. designed and constructed the worm-gear drive to the phase shifter. Of the author's colleagues in the Electricity Division, Mr. E. G. Hope was responsible for the design of the high-frequency generators and all the fixed-frequency receivers, Mr. D. S. Sutcliffe has constructed and maintained much of the other equipment, and Miss R. C. Beeby has taken most of the manual measurements. The development of the monitor proceeded under the general direction of Dr. L. Essen, who, with Dr. L. Hartshorn, advised on the general presentation of the subject-matter.

The work described was carried out as part of the research programme of the National Physical Laboratory, and the paper is published by permission of the Director of the Laboratory.

(6) REFERENCES

(1) Law, H. B.: "An Instrument for Short Period Frequency Comparisons of Great Accuracy," *Journal I.E.E.*, 1947, **94**, Part III, p. 38.
(2) Williams, F. C., and Moody, N. F.: "Ranging Circuits, Linear Time-Base Generators and Associated Circuits," *ibid.*, 1946, **93**, Part IIIA, p. 1188.
(3) Chance, B., *et al.*: "Waveforms" (Radiation Laboratory Series: McGraw-Hill, 1949, Vol. 19).
(4) Drysdale, C. V.: "Alternating-Current Potentiometers and their Applications," *Journal I.E.E.*, 1930, **68**, p. 339.
(5) Mills, J. R., *et al.*: "The Theory and Performance of a High Frequency Current Phase Shifter and its Associated Circuits" (T.R.E. Technical Note, July 1952, No. 158).
(6) Hawkes, H. D.: "Resonant Frequency Meters," *Electrical Review*, 1932, **111**, p. 187.

[The discussion on the above paper will be found on page 173.]

621.396.61 : 621.396.9 : 621.3.018.4(42)

Paper No. 1762 I
Oct. 195

STANDARD FREQUENCY TRANSMISSION EQUIPMENT AT RUGBY RADIO STATIO

By H. B. LAW, B.Sc.Tech., Associate Member.

(The paper was first received 16*th August, and in revised form* 4*th September*, 1954. *It was published in October*, 1954, *and was read befor a joint meeting of the* RADIO *and* MEASUREMENTS SECTIONS 10*th November*, 1954, *and the* NORTH-EASTERN RADIO AND MEASUREMENT GROUP 7*th March*, 1955.*)*

SUMMARY

The MSF standard frequency transmissions take place on 60kc/s, and on 2·5, 5, and 10Mc/s and include modulation frequencies of 1kc/s and 1c/s. The transmissions are controlled by one of three highly stable oscillator-clock chains, which are checked daily in terms of a frequency standard, and are continuously intercompared in order that instability may be quickly detected. Automatic shut-down features are incorporated to reduce the risk of broadcasting incorrect frequencies under fault conditions. The vital parts of the equipment are protected against mains failure.

The transmitted frequencies are normally kept within 1 part in 10^8 of the currently-assessed nominal frequency, based on predicted clock performance, so as to keep within a tolerance of ± 2 parts in 10^8 in terms of finally-corrected time determinations. Day-to-day frequency variations are usually less than ± 2 parts in 10^9.

(1) INTRODUCTION

The importance of standard frequency transmissions as a way of making an accurate standard of frequency widely available was recognized by the International Telecommunications Conference at Atlantic City in 1947. Frequencies of 2·5, 5, 10, 15, 20 and 25Mc/s were allocated for standard frequency services, and the C.C.I.R. was asked to study the problem of providing effective world-wide coverage. Early work, which provided a basis for the Atlantic City decision, and subsequent discussions regarding United Kingdom participation are reviewed in an associated paper;[1] suffice it to say here that the D.S.I.R. asked the Post Office to transmit standard frequencies from Rugby Radio Station. A part-time service was started in February, 1950, and operation for 24 hours per day commenced in May, 1953, on the carrier frequencies 2·5, 5 and 10Mc/s. Transmissions are also made on 60kc/s for one hour per day, from 1429 to 1530 hours G.M.T. The call sign MSF is used.

The conditions at a busy radio station differ greatly from the quiet laboratory atmosphere that is usually considered appropriate for a frequency standards installation, and this fact, coupled with the paramount need to be sure at all times of the accuracy of the transmitted frequencies, gave rise to new requirements, particularly as regards automatic monitoring and fault detection. The present paper describes the installation designed to satisfy these requirements and outlines the methods used to calibrate and control the equipment in terms of primary standards of frequency.

(2) PRELIMINARY CONSIDERATIONS

Accuracy and reliability are perhaps the most important requirements of a broadcast standard frequency service, and the C.C.I.R. has recommended[2] that, at the transmitter, the frequencies should be accurate to within 2 parts in 10^8 of nominal. The day-by-day calibrations of the transmissions have to be made in terms of the unit of time currently available, and appreciable corrections to the initially adopted values may prove necessary when later astronomical data have been taken into account;[3] it is desirable, therefore, for the transmitted frequencies to be held within 1 part in 10^8 of the currently estimated nominal value i order to be reasonably sure of their being within 2 parts in 10 after final corrections have been applied. The frequencies ma be kept in tolerance by frequent small adjustments of the con trolling oscillator, to keep its frequency close to nominal, or b comparatively large adjustments which are made only whe action can no longer be postponed because the working toleranc has been reached. If the controlling oscillator has a good day-to day stability and a low rate of drift, the latter method of adjus ment has the advantage that users interested in the highes precision can readily distinguish between adjustments and drif It was decided to adopt this method of working for the MS transmissions.

There are many things in a frequency-standard oscillator tha can go wrong and affect its accuracy. In an extensive laborator installation the consequences of an oscillator fault are not likel to be serious; other oscillators will be available for checkin purposes so that a user is unlikely to be led astray in his measure ments. Many users of standard frequency transmissions wil however, have no such checking facilities, so it is importan that the transmissions be consistently accurate. Precaution must be taken to avoid the broadcasting of false informatio i.e. inaccurate frequencies, under fault conditions; it is necessar to supplement periodical assessments, in terms of primar frequency standards, with substantially continuous comparison between the frequencies of the controlling oscillator and othe highly stable oscillators. Provided that suitable precautions ar taken—e.g. to avoid coupling between them—stable performanc of one oscillator relative to another suggests that both ar behaving well. The addition of a third oscillator strengthens th argument and allows the identification of one faulty member i the trio; this scheme was adopted for MSF.

The C.C.I.R. has recommended that standard frequenc transmissions include audio-frequency tone modulation and tim signals, the latter to be in the form of 5millisec impulses con sisting of a whole number of cycles of audio-frequency modulatio and repeated at 1-sec intervals, with the 59th impulse in eac minute suppressed. The recommendations leave the choice c the audio modulation frequency, the form of the modulatio cycle and other details to the operating administrations. Dis cussions among the interested parties in the United Kingdon indicated that the most useful modulation frequency was 1kc/ Modulation by the different signals on a consecutive rather tha a concurrent basis was preferred because it allowed a highe percentage modulation by the individual signals. It was als decided that periodical intervals of plain carrier were desirabl to facilitate the use of the carrier frequency in simple receivin equipment. Details of the 15-min modulation cycle eventuall adopted are given in Table 1.

No transmissions are made between 15 and 20min after th hour; this is in accordance with a C.C.I.R. recommendation tha each standard-frequency-transmission station could periodicall interrupt its transmissions to facilitate observations of othe stations and of radio noise. The MSF modulation cycle is quit different from those used at other stations in the standar frequency service, and this helps in its identification.

The requirements as to accuracy and modulation just ou

Mr. Law is at the Post Office Research Station.

Table 1

MSF MODULATION CYCLE

Time				Modulation
Minutes past the hour				
.0– 5		30–35	45–50	1 kc/s modulation
5–10	20–25	35–40	50–55	Time signals
0–14	25–29	40–44	55–59	No modulation
4–15	29–30	44–45	59–60	Morse and speech announcement

:d specify the external performance of the equipment; a ther point that required consideration before the installation ıld be planned in detail arose from its location at a main nsmitting station. Equipment that is to be used at a main nsmitting station must be suitable for operation and main-ance by normal transmitting-station methods. It must be e to run unattended for most of the time, and it must quickly ıw attention to itself when a fault occurs; it must include a r proportion of reserve equipment so that faults can be by-ssed and service restored without waiting for detailed fault arance. Finally the equipment must of course be able to thstand the high ambient fields, the supply-voltage fluctuations d the machinery vibration that are experienced at large nsmitting stations.

3) GENERAL DESCRIPTION OF THE INSTALLATION

The main components of the installation are shown schematically Fig. 1. The equipment divides naturally into three parts: the frequency standard, (*b*) transmitter-drive and announcing uipment, and (*c*) the transmitters. The frequency standard comprises three chains, each formed by an oscillator, a divider and a synchronous clock, and the associated intercomparison equipment that is used to keep a check on stability. Each clock chain produces outputs at 100 kc/s and 10 kc/s, for carrier frequency control, and at 1 kc/s or 1 c/s, assembled in the time sequence given in the previous Section, for modulation; in addition, control signals are produced at the appropriate times for the starting of the announcing equipment and for the 5-min break in transmission once per hour.

Any of the three clock chains, which are designated "red," "white" and "blue," can be selected for the control of the transmitter drive and announcing equipment, which includes 100 kc/s–2·5 Mc/s–5 Mc/s and 10 kc/s–60 kc/s frequency multipliers for the control of the carriers of the short-wave and long-wave transmitters respectively. The announcing equipment automatically produces a Morse and speech announcement when triggered by the "start announcing" signal from the oscillator-clock chain. The announcements and the 1 kc/s and 1 c/s signals are amplified and then used to modulate the transmitters. The short-wave transmitters incorporate their own modulators, but the long-wave transmitter acts as a linear amplifier and is fed with modulated carrier from a modulator in the 10 kc/s–60 kc/s multiplier. All the transmitter drive equipment and the announcing equipment is duplicated.

The h.t. supply to the transmitter-drive equipment is removed between 15 and 20 min after each hour by the "break transmission" signal from the clock chain, so removing the drive from the transmitters. The transmissions are also interrupted by a signal from the alarm system whenever a fault condition implicating the controlling clock chain is detected.

The transmitters are located in the transmitter hall of the Long-Wave Building at Rugby, the low-power equipment being

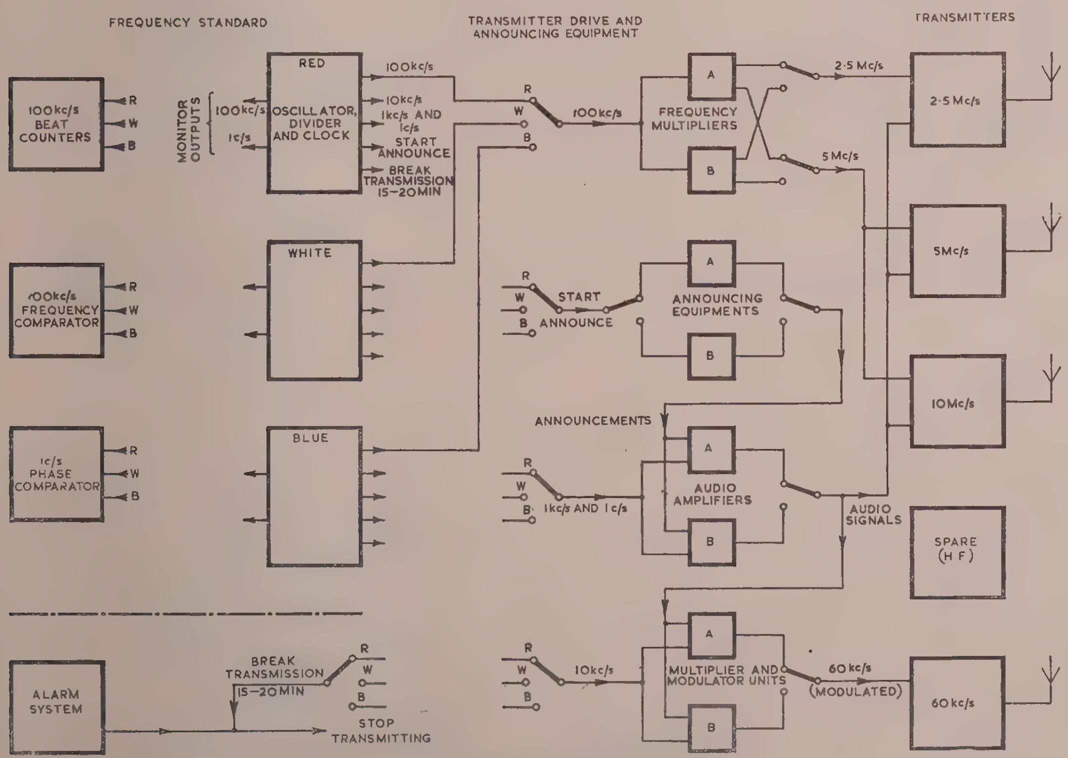


Fig. 1.—MSF Rugby: standard frequency transmission equipment.

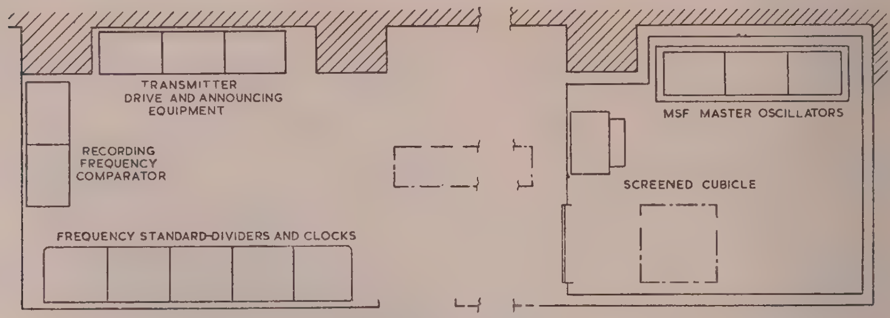


Fig. 2.—Plan of installation.

housed in a screened enclosure (see Fig. 2). The frequency-standard oscillators are additionally protected by a screened cubicle which is thermally insulated and roughly temperature-controlled; the oscillators are mounted on a massive spring-supported plinth to safeguard them from vibration. The enclosure and cubicle were in existence before MSF was planned and they also house some other equipment. Batteries and machines that are used to ensure continuity of operation of the vital parts of the frequency standard during mains failures are housed in other parts of the building.

(4) EQUIPMENT DETAILS

(4.1) Frequency Standard: Clock Chain

Details of the clock chain, made up of an oscillator, a divider and a synchronous clock, are given in Fig. 3. The output of the 100kc/s oscillator is split into four parts, three of which are used externally for carrier-frequency control and monitoring. The fourth output controls a frequency divider giving outputs at 10kc/s (used externally) and 1kc/s, the latter output being applied to a continuously-variable phase-shifter, which also gives four outputs. One of these is amplified and then used to drive the motor of a phonic clock containing a number of mechanically-operated contacts which control the modulation cycle. Another of the phase-shifter outputs controls a frequenc divider to give a 100c/s signal which is then used in a modulato to gate a third output from the phase-shifter in such a way tha the 1kc/s tone is passed and suppressed in alternate 5millise intervals. Contacts in the clock select one of the 100 tone pulse occurring each second and suppress the 59th in each minute, s producing time signals of the type recommended by the C.C.I.R Further clock contacts control the application of these signals t the transmitters according to the MSF programme. The tim signals are also extracted continuously for monitoring purpose and a once-per-minute output is used as well; the frequenc standard monitoring arrangements are described in the nex Section.

The 100kc/s frequency-standard oscillator is of the bridge stabilized crystal-controlled type with resistance-bridge tem perature control for the crystal. It is similar to the Group I type used in the Post Office frequency standard;[4] the cryst oven, which is of a simple form,[5] divides ambient temperatur variations by a factor of about 1 000. Improvements inco porated in the current design include temperature compensatio of the tuned circuits of the oscillation-maintaining amplifi and reduced voltage across the crystal. Also the gas-fille triode in the temperature control unit has been duplicate to reduce the risk of loss of temperature control; the earli

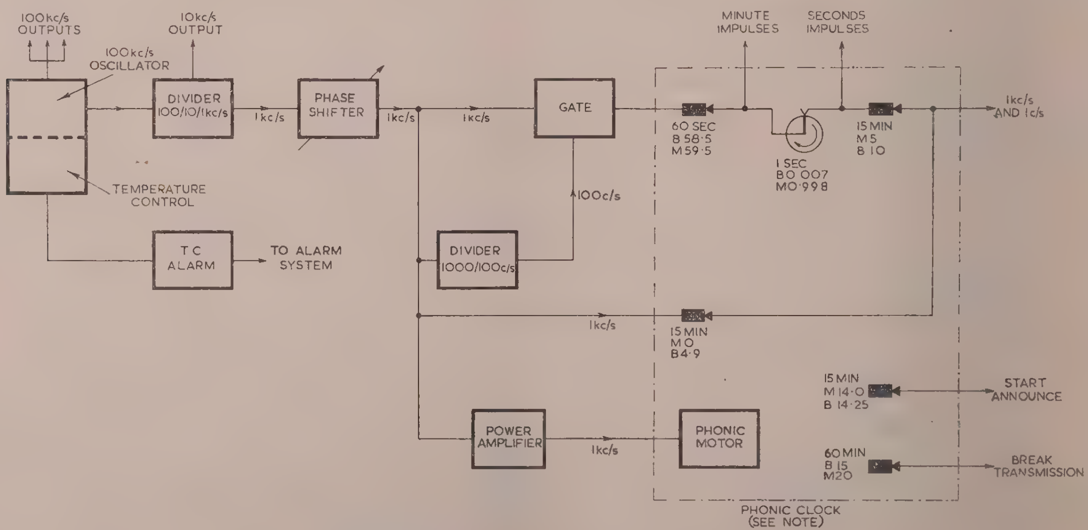


Fig. 3.—MSF Rugby: frequency standard—schematic of clock chain.

The figures associated with the mechanically-operated contacts in the phonic clock show the period of one cycle of operation and the instants during the cycle at which the contact makes (M) and breaks (B).

gle-valve arrangement was liable to sudden and unpre-ctable failures, with disastrous effects on the frequency stability. ith the new arrangement pilot lamps, which indicate the degree load sharing between the two valves, give warning of the terioration of one of them; no harm is done if one of them fails mpletely. If they both fail, or if temperature control ceases r any other reason, an associated unit automatically gives the arm; it operates by detecting the presence of the temperature-ntrol cycle, which normally has a period of about 5sec, and e alarm is given if cycling ceases or if the period becomes nduly long. Two of the three oscillators utilize Essen ring ystals with notably high Q-factors and low rates of frequency ift; advantage is taken of the parabolic frequency/temperature lationship of these crystals by arranging the working tem-erature to be very close to the point where the characteristic as zero slope. Laboratory tests on the completed oscillators owed frequency variations of less than 1 part in 10^9 for supply-oltage changes of 10% or ambient-temperature changes of 10°C; ach condition was maintained for a few hours so that thermal quilibrium was substantially attained. The third oscillator mploys a GT-cut crystal and is not so stable; it is quite adequate or oscillator intercomparison purposes and could be used to ontrol the transmissions if necessary. Information about the erformance of the oscillators in service is given later.

The frequency dividers use a simple form of multivibrator ircuit that has given very satisfactory service in frequency-tandard installations for some eight years. A typical circuit is iven in Fig. 4. Successive stages are substantially identical apart from the values of the anode–grid cross-connecting capacitors, which are chosen to suit the desired natural frequency of oscillation (usually about 60% of the frequency when locked), and the disposition of the interstage locking capacitors, which depends on whether the controlled stage is required to divide by an even or an odd number. The amount of frequency division per stage is modest, division by five being the limit in the MSF equipment. After a divider has been run for two weeks and settled down, the locking capacitors are adjusted for optimum locking, giving correct division over an input frequency range upwards of $\pm 20\%$ relative to the nominal frequency; thereafter it will usually run for several years without attention, and without a slip so long as supplies and the controlling drive are maintained. The small demands made on the valves no doubt contribute to the good long-period performance; the peak anode currents are about 2mA. The dividers are very tolerant of supply-voltage changes; h.t.-voltage variations as large as $\pm 40\%$ merely change the range of input frequency for correct locking by a few per cent and alter the timing of the output waveform by a fraction of a microsecond. Sudden anode-supply-voltage changes are prevented from reaching the multivibrators by the use of a long-time-constant decoupling circuit feeding the common anode supply line of the multivibrators.

The phonic clock, which was designed and manufactured by a contractor, incorporates a commutator turning at 1r.p.s., for defining the seconds impulses, and a number of cam-operated contacts controlled by shafts with revolution periods of 1min, 15min and 1hour. The arrangement is not as simple as Fig. 3 suggests. Where considerable precision of timing is desired in the opening or closing of a circuit, a single contact controlled by a slow-moving shaft is not adequate and it is necessary to supplement the slow-moving contacts with others controlled by faster shafts, which accurately define the instants of opening or closing. The 1kc/s tone and 1c/s impulse circuits are treated in this way, eleven contacts being employed, not including the commutator. The 1min and 15min shafts carry dials.

The remaining units call for little description. The 1kc/s phase-shifter is of the magnetic goniometer type. It is calibrated in fractions of a millisecond and the spindle is geared to a revolution counter. By observing the dial and counter readings while turning the phase-shifter, one may advance or retard the timing of the clock by any desired amount. The gate, which chops the 1kc/s tone into a succession of 5millisec impulses, is a double-balanced modulator using a pair of germanium crystal diodes; it is controlled by a square wave from the 100c/s multivibrator in the 1kc/s–100c/s frequency divider, and the phasing is so arranged that the gate opens just as the instantaneous value of the applied 1kc/s tone is passing through zero.

The frequency-division and phase-shifting equipment of the three clock chains is housed in the three central cabinets of the suite shown in Fig. 5, with the phonic clocks mounted on the top of the cabinets. The sets of equipment are so segregated,

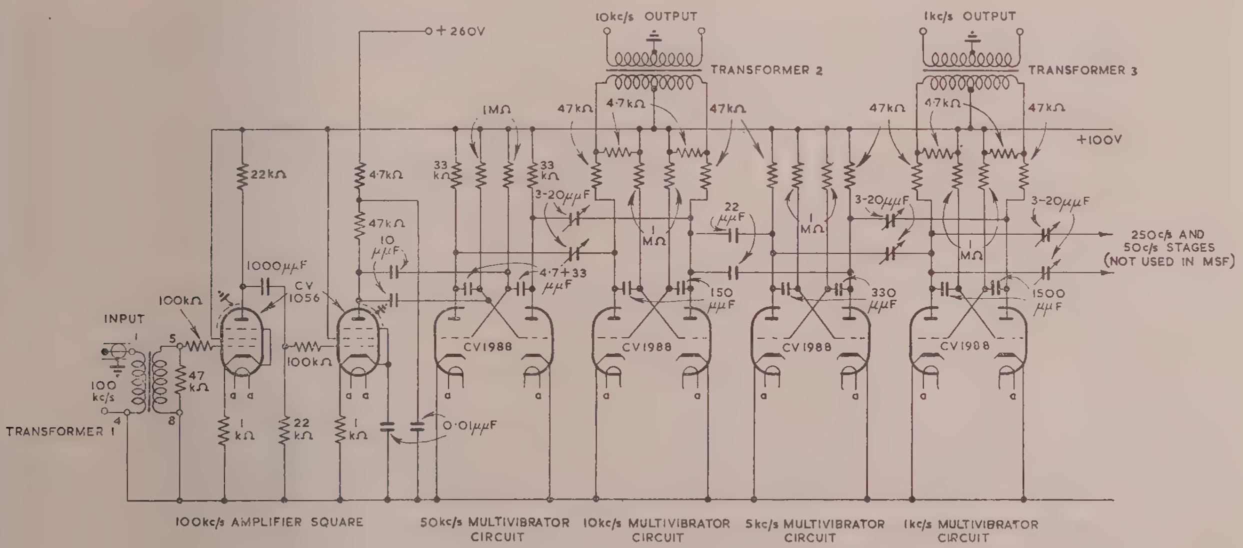


Fig. 4.—Frequency divider 100kc/s–1kc/s.

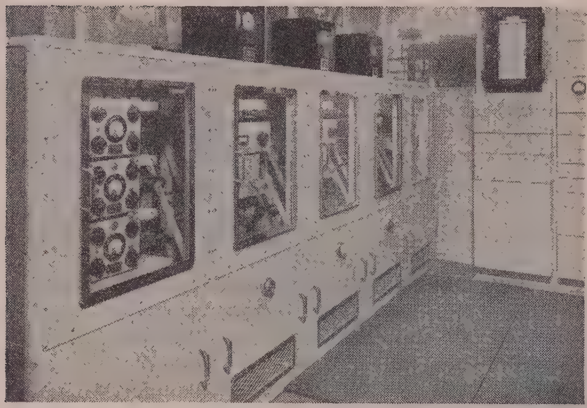


Fig. 5.—MSF Rugby: clock chains, monitoring cabinets and recorder.

physically and electrically, that maintenance work can be undertaken on one without risk of disturbing the others.

(4.2) Frequency Standard: Monitoring Equipment

The frequencies of the three oscillators are periodically compared in a recording frequency comparator, which is shown schematically in Fig. 6. Signals from the pair of oscillators being compared are frequency-multiplied to 5 Mc/s and are then mixed in a ring modulator, the output of which is applied to the coils of a polarized relay. The relay armature oscillates at the beat frequency, which would be 0·1 c/s for an oscillator frequency difference of 2 parts in 10^8. The following circuits are so arranged that each transit of the polarized-relay armature, i.e. each completed half-beat, produces a pulse causing a 50-way uniselector to step. The half-beats are thus counted by the uniselector and this process is allowed to continue for an exact 4 min period. Two of the banks of the uniselector are used, with an auxiliary relay, to form a 100-step potentiometer, and this delivers a direct current proportional to the completed count to a six-colour recording meter. The maximum reading of the system, 99 half-beats in 240 sec, is equivalent to an oscillator frequency difference of just over 4 parts in 10^8; normally the oscillator frequencies should all be within 2 parts in 10^8 of each other, and one bank of the uniselector is used to operate an alarm if this difference is exceeded. The three oscillator pairs, red-white, white-blue and blue-red, are treated in turn, the complete cycle of measurements occupying some 15 min. The 4 min time intervals are defined by minute impulses from one of the phonic clocks, a second uniselector being used to count them. This uniselector controls the selection of the oscillator pairs and is itself controlled by the recording meter to keep the beat-counting process in step with the meter operation; this is necessary because the recorder keeps "mains" time, whereas the beat-counting equipment would keep MSF time if left to itself.

The recording meter is capable of recording a reading every 10 sec, and when it is not actually recording a frequency difference, i.e. for most of the time, it is utilized for recording various supply voltages and the temperature of the oscillator cubicle. The recorder chart speed is low, 2·5 mm per hour, and the record for about six days is visible through the window of the instrument. Any irregularity in the day-to-day operation of the equipment is thus readily observed. The meter can be seen on the right of Fig. 5.

The schematic of a 1 c/s phase comparator is given, with illustrative waveforms, in Fig. 7. The 1 c/s impulses from two of the clocks are amplified and applied to a pair of flip-flop circuits. Each of these is adjusted to fire at the start of the applied impulse, to remain operated for 2·5 millisec and then to restore; a 2·5 millisec positive pulse is produced in the process. The pulses from the two flip-flops are superposed and applied to the grid circuit of a coincidence valve, producing an anode pulse the duration of which depends on the degree of asynchronism of the incoming clock impulses; exact synchronism produces a 2·5 millisec coincidence pulse. The coincidence pulse is integrated giving a voltage proportional to the overlap of the clock impulses. The process is repeated for every pair of impulses from the two clocks; the integrator voltage is allowed to leak away between impulses, the readings being stored in a long-time-constant peak voltmeter calibrated in milliseconds. If the two inputs fall out of step beyond a predetermined limit, normally 1 millisec, an alarm circuit associated with the peak voltmeter comes into action. Thus any appreciable departure from synchronism of the two clocks, whether due to incorrect setting of the phase-shifters or divider slip, or anything else, automatically gives rise to an alarm. Three 1 c/s phase comparators are provided, one each for red-white, white-blue and blue-red comparisons; they are housed in the cabinet at the near end of the suite shown in Fig. 5.

If one oscillator, red say, becomes unstable and drifts away from the nominal frequency, both the red-white and the blue-red excessive-frequency-difference alarms will eventually come into action; this state of affairs permits the automatic identification of the red oscillator as the faulty member of the trio. Similar considerations apply to the loss-of-synchronism alarms. If in these circumstances the red clock chain happens to be the one controlling the transmitters, the transmissions are automatically

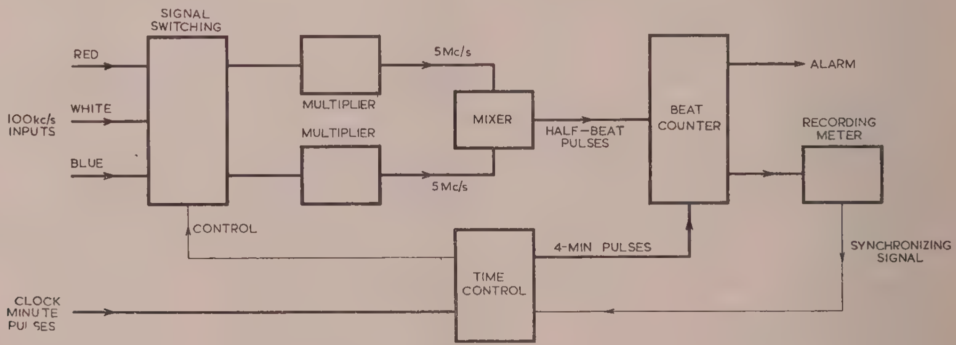


Fig. 6.—Frequency comparator.

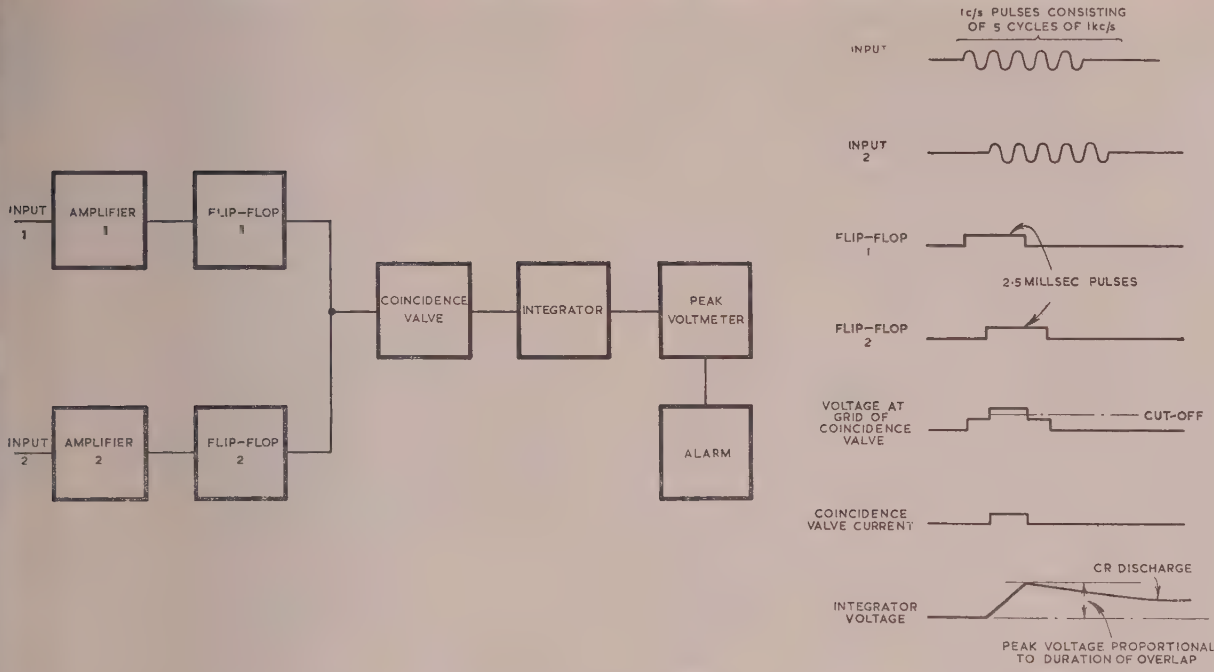


Fig. 7.—1 c/s phase comparator.

nterrupted. The operation of the temperature-control-failure alarm associated with the controlling oscillator similarly interrupts transmissions.

The 100 kc/s beat counters may be regarded as integrating rotary phasemeters. 100 kc/s signals from two oscillators are so mixed as to produce a 3-phase beat-frequency supply, and this is applied to the stator winding of a small motor to give a magnetic field that rotates through one revolution for each beat cycle. The permanently-magnetized rotor tracks the stator field and drives a counter gear train of a type commonly used in electricity meters. There are three of these beat counters, red-white, white-blue and blue-red; they are housed in the far-end cabinet of Fig. 5. Their readings are used in the daily calibration of the equipment, to be discussed in Section 5.

(4.3) Frequency Standard: Protected Power Supplies

Continuity of operation, especially of temperature control, is essential for the best performance of frequency-standard oscillators; it may take weeks for equipment to recover fully from an interruption lasting only an hour. The MSF equipment therefore incorporates a protected power supply. In addition to the oscillators, the frequency dividers and the clocks benefit from this supply, as do the beat counters, to preserve the continuity of long-period calibration data and to avoid the need for resetting the clock timing. The 1 c/s phase comparators also receive protected power supplies.

The protected h.t. supply is derived from a 240 volt 50 Ah battery that operates on a float-charge basis; the immediate effect of a mains failure is merely a slight fall in voltage, which is of no consequence. The protected a.c. supply for valve heaters and crystal ovens is normally taken direct from the mains, but it is automatically transferred to a motor-alternator, which draws its energy from the battery, if the mains voltage varies more than $\pm 10\%$ relative to nominal. An alarm is given; resetting is a manual process. The motor-alternator runs continuously so that the interruption to the a.c. supply on change-over, determined by the operating times of the control gear and contactors, lasts less than a second. The thermal inertia of the valve cathodes and crystal ovens renders the disturbance insignificant. Duplicate batteries and motor-alternators are provided, the battery change-over switch being of the make-before-break type.

(4.4) Announcing Equipment

The announcing equipment was designed and made by a contractor. The Morse and speech announcement, of about 1 min duration, is derived from a magnetic-tape record in the form of a continuous 50 ft loop, the bulk of which is housed, as a loosely wound helix, in a cassette. In between announcements the equipment rests with its valves "hot"; on receipt of a "start announcing" signal the tape-deck motor is switched on and the tape is drawn past the reproducing head. The process continues until the complete loop has been drawn through, when the motor is automatically stopped by a pair of contacts detecting a short length of tape marked with conducting paint.

Recording facilities are provided so that new tapes can be prepared on the spot with a minimum of delay.

(4.5) Transmitter Drive Equipment

The circuits of the transmitter drive equipment are conventional and call for little comment. The frequency multipliers are of the class-C type, the maximum multiplication per stage being 5 times; the h.t. supply to the 10–60 kc/s frequency-multiplying stages is stabilized in the interests of phase stability. Measuring equipment for the checking of the carrier and audio feeds to the transmitters is provided, and there is a loudspeaker for monitoring the outgoing audio signals. The h.t. power supplies to the transmitter drive panels are automatically interrupted at 15–20 min past each hour, under the control of clock signals, and whenever necessary under fault conditions, so causing the transmissions to cease.

(4.6) Transmitters

The short-wave transmitters are of a standard commercial double-sideband type. Each transmitter is pretuned for operation on 2·5, 5, 10, 15 and 20 Mc/s and has a carrier power of about 500 watts; at present the four transmitters are used in rotation to provide a 24-hour service on three frequencies, 2·5, 5 and 10 Mc/s, one transmitter always being available for maintenance. A bottom-fed mast radiator is used on 2·5 Mc/s, and the aerials for the other frequencies are quadrant dipoles.

The long-wave transatlantic telephone transmitter, which is not often needed for traffic in the afternoon, is used for the 60 kc/s service; a carrier power of about 10 kW is delivered to the 800 ft aerial.

(5) OPERATION AND PERFORMANCE

The carrier frequency of the 60 kc/s transmission, which is comparatively little affected by Doppler variations due to changing ionospheric conditions,[1] is measured daily, except on Saturdays and Sundays, in terms of the frequency standard at the Post Office Research Station, Dollis Hill. The measurements are extended over a period of several minutes to give a comparison accuracy better than 1 part in 10^9. Measurements are also made by the Royal Observatory and the National Physical Laboratory. An assessment of the absolute frequency of the oscillator controlling the transmissions is thus obtained. The mean frequency differences between this oscillator and the other oscillators at Rugby over the previous 24 hours is calculated from the beat-counter readings, and so the absolute assessment is extended to the reserve oscillators. In the light of these assessments the National Physical Laboratory in consultation with the Post Office decides whether any frequency adjustments are necessary to keep the oscillators within 1 part in 10^8 of nominal frequency. As has already been indicated, it is the policy to adjust the controlling oscillator as infrequently as possible; for this reason the working tolerance is slightly elastic in practice.

The timing of the 1 c/s pulse signals is also measured regularly. As these signals are intended primarily for frequency calibration, quite large time errors are allowed; when necessary, corrections are made in steps of 50 millisec on the first day of the month and a statement about the correction is included in the announcement. The time intervals defined by the 1 c/s signals thus have the same accuracy, at the transmitter, as the carrier and audio frequencies. The phases of the reserve clocks are adjusted daily so that all three clocks will remain within ± 1 millisec of each other during the ensuing 24 hours. Thus, if the controlling clock should fail, the service can be restored quickly, using a reserve clock, with very slight disturbance to the timing of the 1 c/s signals.

Fig. 8 shows the results, over a typical two-month period, of the current-frequency assessing process just described. The upper three curves are assessments of the three oscillators. During the period under review the transmissions were under the control of the blue oscillator, which had last been adjusted over a year previously (6th May, 1953); its curve gives a good idea of the day-to-day consistency of the 60 kc/s carrier-frequency measurements. The lowest curve shows the daily 24-hour mean-frequency differences, calculated from the beat counter readings, between the red and blue oscillators, these being the two with Essen-ring crystals. The performance of the white oscillator is not so impressive, although it is adequate for a second reserve. The rather rapid drift in the early part of June arose from a power disturbance on the 3rd June, the first occasion since the completion of the equipment on which the protected supply failed; the effect of this disturbance on red and blue was very slight. The small jump in the measured frequencies on the 15th July arose from the method of calibration of the Dollis Hill standard. Current assessments are based on monthly predictions by the Royal

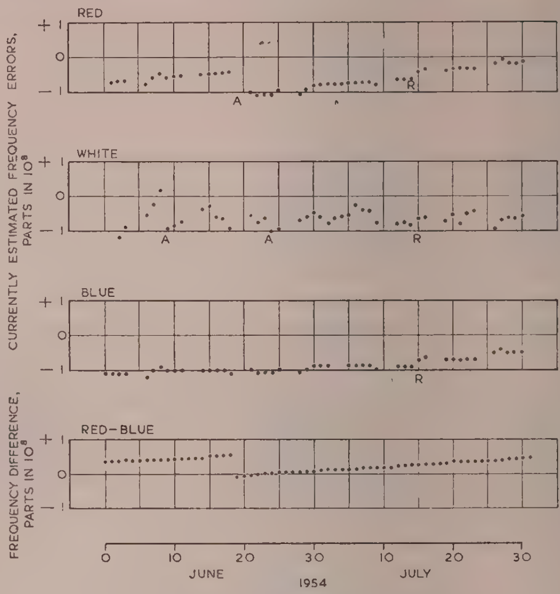


Fig. 8.—Oscillator performance.
A = Oscillator adjusted. R = Standard reassessed.

Observatory of the frequency of an oscillator in the standard fc the ensuing month; successive predictions are in general nc continuous one with another, so that current calibrations are liabl to show a discontinuity at mid-month, when the new predictio is received.

Frequency comparator records show that variations of th frequency difference between blue and red rarely exceed ± 2 par in 10^9 during a 24-hour period; analysis of 92 comparison recorded in a typical 24-hour period (13th–14th June, 1954) gav a standard deviation of 7·6 parts in 10^{10}. The short-period scatte of comparisons involving the white oscillator is slightly greate and there is a tendency for irregular drifts, which last a few hour giving a total range of variation of up to ± 3 parts in 10^9 in 24-hour period.

The long-term frequency accuracy of the MSF signals, in term of finally-corrected time, is discussed in another paper,[1] to whic reference may also be made for time-signal results.

(6) FUTURE DEVELOPMENTS

A feature of standard-frequency-transmission work during th last few years has been the steadily increasing importance c the time signals, and this has led to various proposals for the improvement. The present form of the signals has been criticizec on two main grounds, first that the energy content is low so tha the signals are vulnerable to interference, and secondly that th "negative" minute signal—the 59th-second impulse is omitted— is particularly hard to recognize under difficult conditions. Th MSF equipment is being rearranged at the request of the Nationa Physical Laboratory to produce a positive minute signal in th form of a lengthened impulse at the 60th second. Also there a proposal for the time signals to be radiated for nine minutes i each quarter-hour period, from 5 to 14 min after the exac quarter, instead of for five minutes as at present; the loss of th carrier-only interval that would result from the propose rearrangement is unimportant, for experience has shown that th presence of the time signals does not appreciably hinder the us of the carrier for frequency calibrations. There is not at preser any plan to increase the power of the MSF time-signal impulses

y substantial increase in the power would require transmitters high-peak-power rating, perhaps of the pulse type, involving nsiderable extra expense.

Experience has shown that the 60 kc/s transmission is far perior to the short-wave ones for accurate work, so that the estion arises whether further experiments with long-wave mals may be desirable. In this respect the location of the MSF quency-standard equipment at a main transmitting station ovides interesting possibilities; the equipment is in fact already ing used to control the carrier frequency of the high-power kc/s telegraph transmitter GBR.

(7) CONCLUSION

The results given by MSF show that quartz-clock equipment the highest accuracy can be operated successfully under radio-ation conditions; it may be claimed that this type of equipment s graduated from the laboratory. The accuracy of the trans-issions is at present limited by the uncertainties of absolute equency determination. The day-to-day frequency variations the equipment are about ± 1 part in 10^9, and it is to be hoped at work on frequency standards utilizing atomic resonances ay soon allow the equipment to be calibrated with the accuracy at its stability merits.

(8) ACKNOWLEDGMENTS

The MSF equipment is the outcome of team work by many dividuals in the Radio and Research Branches of the Post Office Engineering Department; particular credit is due to the Dollis Hill crystal group, under Mr. J. E. Thwaites, and to Messrs. A. K. Dobbie, J. E. Haworth and J. S. McClements, who were responsible for most of the detailed design work. The successful operation of the service rests on co-operation between the Royal Greenwich Observatory, the National Physical Laboratory and the Post Office, and on the careful work of the staff at Rugby Radio Station. The author wishes to express his thanks to Mr. H. T. Mitchell, for his interest and encouragement, and to the Engineer in Chief of the Post Office, for permission to publish this paper.

(9) REFERENCES

(1) Essen, L.: "Standard Frequency Transmissions," *Proceedings I.E.E.*, Paper No. 1677 R, July, 1954 (**101**, Part III, p. 249).
(2) C.C.I.R. Recommendations: No. 18 (Stockholm, 1948), No. 70 (Geneva, 1951) and No. 122 (London, 1953).
(3) Smith, H. M.: "The Determination of Time and Frequency," *Proceedings I.E.E.*, Paper No. 1054 M, November, 1950 (**98**, Part II, p. 143).
(4) Booth, C. F., and Laver, F. J. M.: "A Standard of Frequency and its Applications," *Journal I.E.E.*, 1946, **93**, Part III, p. 223.
(5) Creighton, J. L., Law, H. B., and Turner, R. J.: "Crystal Oscillators and their Application to Radio Transmitter Control," *ibid.*, 1947, **94**, Part IIIA, p. 331.
(6) C.C.I.R. Report No. 29 (London, 1953).

DISCUSSION ON THE ABOVE TWO PAPERS, AND A PAPER ON "STANDARD FREQUENCY TRANSMISSIONS,"* BEFORE THE RADIO AND MEASUREMENTS SECTIONS, 10TH NOVEMBER, 1954

Dr. R. L. Smith-Rose: I recall that we started out with the ransmitter at the National Physical Laboratory with the idea of adiating regularly programmes involving 16 separate frequencies. These were not sent out simultaneously, but in two batches of ight on one day in the month. It is evident that this was too mbitious a programme, which even the Americans have never ried to imitate, and we soon abandoned it for the two separate requencies described in Dr. Essen's paper.

The war broke the continuity of the development of this type f service. It is now nearly ten years since I became interested n ensuring that a suitable and up-to-date service of standard requency transmissions was set up in this country. I felt that it vas necessary on several grounds. First, I wanted to see us ake our place in the forefront of the field in this subject—a osition which we had been in danger of losing to the Americans. econdly, I foresaw an international situation in which we were oing to be asked how to ensure world-wide coverage for a tandard frequency service, and obviously this had to be developed n the lines adopted by the United States Bureau of Standards. This "got off to a very quick start" at the end of the war by adiating on all six frequencies which had been laid down by the Atlantic City Convention for this type of service. I was hopeful hat we could start a service, perhaps on a more modest basis, ut at any rate on these lines.

There were several difficulties in those early days, some tech-ical and some of an entirely different character. For example, here were three Government Departments interested in this atter with somewhat different points of view on the question of vhose responsibility it was to organize and operate the service.

After some discussion with representatives of the Royal Greenwich Observatory and the Post Office it was decided that he National Physical Laboratory, through the Department of cientific and Industrial Research, should pay for the service, and ave complete technical responsibility. The service developed on those lines. The National Physical Laboratory pays the Post Office for all the elaborate equipment which has been built and used at the Rugby station.

Another matter which was discussed in some detail was whether we could start the service in this country on the same frequencies as those used in America, but arranging the schedule of transmissions in such a way that any interference between the transmissions from the two stations could be reduced to a minimum. We spent some time studying charts of ionospheric propagation time in order to evolve a schedule whereby certain frequencies were used at certain times of the day and seasons of the year. However, we soon came to the conclusion that if we adopted any such schedule it would be so complicated that no prospective user would ever be sure what any one of the frequencies was at the time when he wanted to use it, and so we abandoned this scheme and decided that if this question of mutual interference did come to a head—and it will do so in the future—a simple division of time would have to be adopted for the programmes of transmissions from different countries.

Having settled the question of who was going to pay, and that there was nothing magic about the choice of the time, things began to move rapidly and the service developed on the lines described in the papers. In addition to establishing the service and making it available to all users, it is necessary to tell the world what has been accomplished. These three papers present the whole story of this technical achievement. Dr. Essen's paper describes the development of the service from the earliest days and gives some account of its present scope. Mr. Law's paper presents an excellent account of the special equipment developed, built and installed at the Rugby station, and Mr. Steele's paper describes the monitoring equipment which has been set up by the National Physical Laboratory to keep a watchful eye on the frequencies as they are transmitted in these programmes.

Dr. Essen stated that he had had a large number of replies from

* Essen, L.: Paper No. 1677 R, July, 1954 (101, Part III, p. 249).

people who had received these transmissions. Can he discriminate between the people who merely send a card saying that they have received the transmission and those who are able to make use of the transmissions for their proper purpose? When one is asked who uses this service and how one can justify the expenditure of the money necessary to provide it, it is not easy to reply.

Mr. Law describes an automatic arrangement whereby, when the "red" clock chain frequency departs from a certain tolerance, automatic apparatus comes into operation and the service of transmissions is interrupted. I believe that a similar interruption takes place when the temperature gets out of control in some parts of the equipment. How many times has this interruption taken place since stage 2 of the service was started 18 months ago, and what is the total service time which has been lost owing to these causes?

My last point concerns the use of the frequency of 60kc/s for transmission. As will be generally realized, this is not one of the frequencies which was allocated internationally for this type of service, and Mr. Law has explained that this frequency, which was allocated to the Post Office for a radiotelephone service, has been used for this other purpose because the telephone service happens to be slack or non-existent in the afternoon. However, what happens when the traffic on the radiotelephone service becomes heavy? Are the standard frequency transmissions interrupted in order to put the telephone calls through, and if so, does the Post Office make a rebate to the National Physical Laboratory in order to cover the revenue obtained from these telephone calls?

Mr. H. M. Smith: I should like to make a few remarks about the two functions performed by the Royal Greenwich Observatory in connection with the transmissions. The first is concerned with the establishment of the fundamental time required *inter alia* for the calibration of the frequency and time intervals, and the second with the routine measurement of the time and frequency of the transmitted signals.

The unit of time is at present defined in terms of the rotation of the earth by means of astronomical observations. There are, however, certain variations in the rate of rotation which can no longer be ignored, and it has become necessary for the Royal Observatory to provide a time system which is effectively uniform while remaining based on the rotation of the earth over periods of one or two years. This provisional uniform time system is the basis of the monthly estimates made by the Royal Observatory of the frequency of selected standards.

This procedure for the establishment of a uniform time system is limited by the period of uniform run of high-precision clocks. In astronomical investigations of the motions of the planets it is necessary to employ a uniform time scale extending over several centuries, and a new time system for this purpose has recently been adopted.* This is known as Ephemeris Time, and is based on the revolution of the earth around the sun, i.e. on the year rather than the day; it is thus free from the effects of irregular changes in the rotation of the earth. In practice it is measured by observations of the position of the moon. A new international programme of photographic observations of the moon against the background of the stars is designed to determine Ephemeris Time with so little delay that it will be available as a basis for current frequency estimation.†

In the paper by Mr. Law mention is made of the possibility of a fundamental standard based on atomic or molecular resonance. Such a standard would provide only frequency and time intervals, whereas Ephemeris Time offers a true time system, and already has the advantage of the link of nearly three centurie with Greenwich Mean Time. Atomic frequency standards wi serve to provide a convenient link between the long-term funda mental astronomical time standard and the current calibration o frequency standards to high accuracy.

The Royal Observatory maintains a regular check on man British and foreign time signals and standard frequency emissions Frequency comparisons on the 60kc/s MSF transmission ar made with two independent Observatory standards, and check on the accuracy of measurement is maintained by com paring the frequency difference between the standards measure directly and via the MSF transmission. Over a recent perio (1st July to 8th November, 1954) the discrepancy was withi $\pm 5 \times 10^{-11}$ for 98% of the measurements made. For th superimposed one-second "ticks," the total scatter of the indi vidual readings on a given day was within ± 15 microsec for 92% of the measurements. The day-to-day accuracy of compariso is influenced by propagation conditions, but remains full adequate for all practical purposes.

Mr. R. S. J. Spilsbury: The technical and financial responsi bility for this work lies entirely with the National Physica Laboratory, but it is obvious that we should be incapable o carrying it out without the assistance of various people. W cannot do this job without time, and the Royal Greenwic Observatory provides us at present with the fundamental data Of course, the earth goes round in rather a "wobbly" way an the stars are usually invisible in this country, but we are a present making an atomic clock which will avoid these effects.

Dr. Essen was one of the prime movers in this work becaus 17 or 18 years ago he invented the Essen ring which oscillates in a peculiar way without changing the length of its perimeter, and I would claim that this is the most stable standard of any physica quantity which exists to-day. It has been developed in a portabl form by the Post Office, who have devised a mounting using sil strings as supports.

The Post Office are responsible for the transmissions—in fac the National Physical Laboratory is only a watchdog in th background. However, whilst the Post Office do this wor extraordinarily well, they do not, of course, do it for nothing Their charges plus the expenses of monitoring, etc., at th National Physical Laboratory amount to nearly £10 000 a year

Dr. Smith-Rose asked who uses this service and what the use it for. It would be a great help if any people who use thi service and find it valuable would let us know. This is, of course a public service, and we should like it to be as good as possible For example, I am being pressed by one colleague at the momen to change over from 1 000c/s to 440c/s for the a.f. modulation on the ground that it would provide the British Standard pitc for musicians (actually it is about 100 000 times more accurat than is necessary for their purposes). For scientific purpose 1 000c/s is more convenient, and 440c/s more difficult to conve into anything useful. It would be interesting to know whic frequency is preferred. It is extraordinarily difficult to get suc information.

Mr. N. Lea: I have recently travelled to the United States an seen the WWV installation. I visited the National Bureau o Standards, which controls WWV, and the Naval Observator which provides the time information; and I think that in thi country we have an arrangement which compares favourably.

According to Dr. Essen's paper, MSF gives a reliable 1-hou signal which can be received with an accuracy of within ± 2 in 10^9—I use the word "accuracy" with some qualification—s that for one hour a day we have an accuracy of within ± 2 in 10^9, while for 24 hours a day we have an accuracy of withi about ± 2 in 10^7. This means that, so long as it has at it command a number of well-tested oscillators, an industrial labora-

* SADLER, D. H.: "Ephemeris Time," *Occasional Notes of the Royal Astronomical Society*, 1954, 17, p. 103.

† MARCOWITZ, W.: "Photographic Determination of the Moon's Position, and Applications to the Measure of Time, Rotation of the Earth, and Geodesy," *Astronomical Journal*, 1954, 59, p. 69.

ry is in a position to make frequency measurements at any time f the day to a resolution of 10^{-8} and to be engaged in the development of stable oscillators to about the same stability. If we want develop oscillators which are stable to within 1 part in 10^9 for seful periods, it means that the MSF service, although a very ood starting point, is not entirely adequate, and we must set up cally oscillators which are stable for at least 24 hours to a esolution of better than 10^{-9}. In using the MSF transmissions this way for the development of stable oscillators, we must, f course, be very careful about the definitions of frequency and hat is being measured, and we must clearly distinguish between esolution and accuracy. It is particularly important that we hould always attach to any statement of a frequency comparison e period over which this comparison was made. Many of the omparisons mentioned in the papers are concerned with periods f observation of a few seconds only. For general intercomarisons between laboratories, and even between oscillators in he same laboratory, I think that this is unnecessarily and perhaps angerously short. I prefer a period of 200 sec in order to inimize uncertainties owing to effects in the comparison pparatus itself.

With regard to the assessment of oscillator performance, even hen extensive recorded information is available it may be ifficult to predict future performance with a desired degree of robability. About three years ago I obtained recordings which ndicated inter-oscillator stabilities of 10^{-9} continuously over a eriod of 30 days (Fig. A). This seemed a good result and I felt at the time that some use could be made of it. However, I soon discovered that various erratic effects associated with the electronic system were quite unpredictable, and it has taken about wo years to find a solution to the difficulty. I give this history of my own misfortune as a warning to others working in this field.

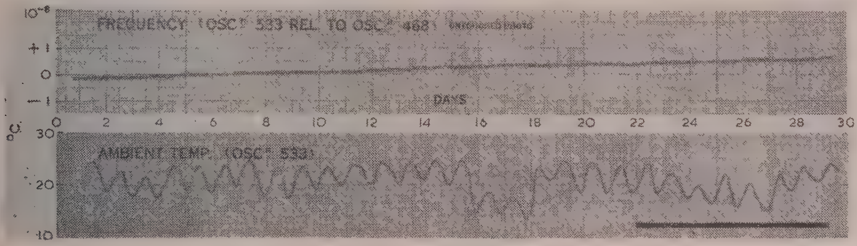


Fig. A.—700-hour automatic record.
More than 99% of 4 500 plotted points are within 1 in 10^9 of a linear drift.

Fig. 8 of Mr. Law's paper shows an intercomparison between ring oscillators, and there is a scatter which seems to be rather greater than 1 or 2 parts in 10^9. I wonder whether this is due to actual frequency differences between the oscillators or whether it can be ascribed to some irregularity in the recording arrangements.

In Fig. 4 of Mr. Steele's paper there is a slight scatter in the difference between the ring oscillators, and it looks as if this is also about 10^{-9}. Will the author confirm whether this is so, or whether it is less than that?

Capt. C. F. Booth: Apart from particular applications requiring very high stabilities, i.e. navigational aids, common-frequency working in broadcasting, etc., we must make more effective use of the spectrum by improving the frequency tolerance of transmitters if we are to accommodate the ever-growing needs. It is essential, therefore, for the authors' work to go forward and for those concerned with the planning, provision and operation of radio services to use the information they provide. My own experience has shown that the stability achieved in the laboratory is quickly applied at radio transmitting stations. I would emphasize that the elegant frequency standardizing equipment at the Rugby Radio Station described by Mr. Law is operated by personnel normally concerned with the maintenance of transmitters for long-distance communication.

I would like to deal with two points about the programme of future work. The first is the development of a fundamental frequency standard which is independent of the vagaries of the quartz-crystal oscillator; I hope that the work being carried out by Dr. Essen on the use of atomic resonance will go forward quickly. For the radio-station operator, the authors should seek means of achieving the very high stabilities they refer to, with relatively simple equipment; this will call for a further reduction in the ageing of the quartz crystal element, and a simplification of the crystal mounting.

Mr. H. V. Griffiths: For many years we have rated our frequency standards at the B.B.C. measuring station against GBR time signals broadcast by the Post Office and controlled by the Royal Greenwich Observatory, but we have found the MSF 60 kc/s transmissions very useful as a frequency reference with which we can conveniently compare our own standards.

In the introduction to Dr. Essen's paper it is suggested that television will increase the stringency of requirements for the closeness of control of transmitting stations. That may not apply, because at the moment none of these stations are synchronized. They are offset by a deliberate fraction of the line frequency. The stringency of requirements is sufficiently obvious in the medium-frequency band, because if we have to synchronize, or approximately synchronize, two transmitters, e.g. at a frequency of 1 000 kc/s, then assuming that quite a low tolerance is adopted, say 5 sec, for the beat frequency between them, it is necessary to have the transmitters maintaining their allocated frequency to within 0·1 c/s, and that is 1 part in 10^7. The measuring station, of course, must be one order better, which brings us to 1 part in 10^8. Thus the need for these accuracies is well established.

The 200 kc/s Droitwich transmission is greatly used in the United Kingdom, and to some extent on the Continent, as a substandard of frequency for quasi-scientific purposes. I believe that the reason for this is that it is available almost continuously, whereas the MSF 60 kc/s transmission, in the most stable propagation frequency band, is only available for one hour a day. There is undoubtedly a need for some long-period low-frequency transmission in the 50–200 kc/s band, where propagation is most stable.

In Section 5.1 of Dr. Essen's paper there is mention of the scatter of the results. In making comparisons with MSF we obtained apparent scatters of the order of 4 or 5 parts in 10^9, but we decided that part of the scatter was a systematic error in the readings obtained from the apparatus used for intercomparing MSF and Tatsfield. The apparatus we use is rather less complex and costly than that described in the papers. It relies partly upon manual operation. The MSF 60 kc/s transmission and the

local standard are displayed together on a circular oscillograph trace, the phase alteration being determined over a period of 500 sec. Another set of apparatus permits the intercomparison of our own standards at a frequency of 20 Mc/s, by multiplication of the fundamental 100 kc/s frequency of the standards.

Fig. 2 of Dr. Essen's paper shows the fluctuations in received frequency with the MSF short-wave transmission. Similar fluctuations are observed in Tatsfield measurements of the B.B.C. substandard frequency transmissions on 17 810, 9 510 and 6 180 kc/s. We notice the same sunrise and sunset effects, and the diurnal changes of sign of the error frequencies.

With regard to Mr. Law's paper, I would be interested to have precise details of the device which detects the presence or absence of the temperature-control cycle and operates the alarms. There are existing alarm systems of Post Office design on the Tatsfield quartz ring and on other B.B.C. standards, but the particular device described may be new. The No. 4 ring standard, which was put into service in November, 1953, has been very satisfactory indeed up to the present. Its rate has in the last few months become asymptotic to about ± 4 parts in 10^8 and it is extremely stable to within limits of 1 or 2 parts in 10^9, judged from the intercomparisons with MSF, in which the errors of propagation and intercomparison may not be negligible.

I notice from Fig. 4 of Mr. Law's paper that multivibrators have been adopted for use at Rugby. The choice may be understandable with so many circuits to control. At Tatsfield, on the other hand, we prefer the so-called "knock-back" (regenerative) binary circuit for frequency division, and we have two systems in operation, each frequency-dividing from 100 kc/s to 1 c/s. The 1 sec pulses are not only used for time comparisons with GBR but have been made to drive the station clocks provided for ordinary operational purposes. We are very satisfied with these modified binary divider circuits.

I notice that the 59 sec impulse from MSF is in process of being changed. I think there will be general agreement that the lengthening of the minute dash is to be preferred. Otherwise, one can never be quite sure whether the signal has faded out or not.

We have nothing as complex or elegant as the apparatus described in Mr. Steele's paper. However, for the very large number of measurements that we make, the B.B.C. comparison equipment gives results that compare very well with the MSF curves shown. Our results prove that the Tatsfield No. 4 standard has a performance about equal to that of the MSF standard.

With regard to Fig. 5 of Mr. Steele's paper, I assume that the conversion standard at the bottom left-hand end of the diagram is the same as that which is used to compare with MSF. In the N.P.L. measurement of 200 kc/s I do not see any reference to the small phase changes that seem to occur with peak modulation. They are very small—about 1° or less, 1° being about the maximum. We have observed this effect with our apparatus. One should always, of course, suspect trouble in the comparator circuits, but when a local standard is substituted for the Droitwich carrier, the phase shifts do not occur.

Mr. H. T. Mitchell: I feel that the authors have been too reticent about the ring itself, which is the heart of the equipment. Dr. Essen described the Z-cut ring in 1938. It was a plain ring of quartz mounted on three pins, with the electrode system spaced from the quartz by a small air-gap. Since 1938 the Post Office have taken that basic ring and put the electrodes on the quartz surface by evaporating silver, and mounted it not on three pins but in six small silk threads. I think that Dr. Essen is of the opinion that we have not made it any more stable, and he may be justified in that view, but at least we have made it transportable. We can now make a ring in the laboratory, age it and send it several thousands of miles, being fairly certain that it will give as good a performance when it arrives at the user's premises as it did when it left us. This is easily the best resonator available to-day.

Quite a number of these thread-suspended rings have been made, and 38 are in use at the present time—three in Australia, five in the United States, four in Canada, and the remainder in this country. Of these latter, six are at the Royal Observatory at Abinger, three are at the National Physical Laboratory, one is at the B.B.C., and the remaining 16 are in use by the Engineering Department of the Post Office. That may seem a very large number to have at the Post Office, but they are at various places, and some are used as oscillators and some as resonators.

Fig. B is a photograph of this ring showing the plated quartz ring supported on fine silk threads.

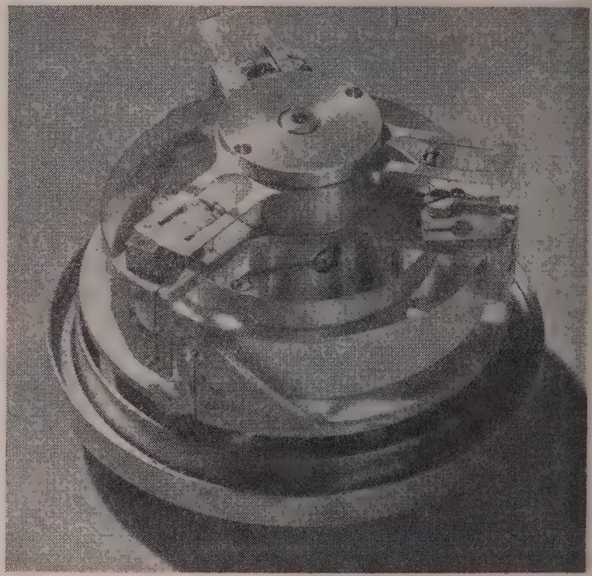


Fig. B.—Quartz-crystal oscillator unit W6.

100 kc/s Z-cut quartz ring mounted on thread suspension in crystal holder W6, with cover removed.

Mr. G. B. Wellgate: The equipment at the Royal Greenwich Observatory, dealing with the frequency comparison between the MSF carrier of 60 kc/s and the local standards, consists of a signal amplifier incorporating a crystal-lattice filter of about 1 c/s bandwidth followed by a convertor from 60 to 100 kc/s.

This convertor is essentially a regenerative divider by a factor of three. In such a divider a frequency of 40 kc/s derived from a 20 kc/s signal is used to modulate the input signal thus producing two sidebands—the 20 kc/s signal mentioned and one of 100 kc/s wanted for comparison purposes.

Measurements made at quarter-hour intervals result in frequency differences between the carrier of MSF and the local standards integrated over three-quarters of an hour. The mean value for the anomaly mentioned by Mr. Smith is just over one part in 10^{11}, and the measured differences between the carrier frequency of MSF and that of the best local oscillator do not normally depart by more than ± 3 parts in 10^{10} from a linear drift.

The method of measurement is due to R. H. Tucker of the Royal Greenwich Observatory.

The time-signal receiver for 60 kc/s has an effective bandwidth of 4·6 kc/s, and mean deviation of individual measurements departs by ± 4 microsec from the mean of a series of such measurements.

Mr. P. P. Eckersley: It is clear that the frequency of a crystal

llator can be maintained to within very close limits; as a llary there obviously exists a means whereby time standards d be established to a high degree of accuracy.

he authors have dealt with international checks upon frency and how they are made; I would like to deal with the sibilities of establishing accurate time nationally and in a nner so that all could have an easy reference to it. More cisely, could we not assume that public and private clocks ld be maintained to the same standard, and that an accurate ?

Ve can hear, in our own homes, the strokes of Big Ben, the adcast pips, and if we have a telephone, the TIM service. en we go out of doors it is another matter; we have to rely on the vagaries of a watch or public timepieces.

What are the authors' views on the possibility of synchronizing mains frequencies with the submultiple frequency of a crystal oscillator. The crystal-oscillator frequency could be divided to become 50c/s, and this could be compared with the mains frequency. The out-of-synchronism component could then be used to open and shut the steam valves at the generating station. The problem is well known and concerns that of servo mechanisms; its application, to say the least of it, would be on a rather large scale. The fact that the scale is large does not, however, vitiate a principle, and some superficial consideration of the matter leaves me with the feeling that, if it were worth doing, it could be done.

Assuming that the scheme were successfully consummated, we would banish once and for all one of the annoyances of life—the inaccuracy of public and domestic clocks.

I understand that in the United States an accuracy to within 100 millisec for a mains-driven clock is guaranteed.

THE AUTHORS' REPLIES TO THE ABOVE DISCUSSION

Dr. L. Essen (*in reply*): In reply to Dr. Smith-Rose I can give ew figures about the reception reports we have received. The jority (450) come from amateur radio operators, and of these, me only required an acknowledgment card while others stated t they made use of the transmissions for calibration purposes. l these reports give useful information about the reception ea. The number of commercial firms and Service and unirsity laboratories in this country known to make use of the nsmissions is approximately 150, but we suspect that the actual mber of users is considerably greater, because they do not, in neral, write to us unless they have some query to raise. In dition there are at least 15 specialist users, including several servatories and firms developing high-precision oscillators, ho make use of the full accuracy of the low-frequency transissions. In this connection I might mention that the work at arvard University, to which brief reference is made in the paper, s continued, and that to assist such long-distance reception e Post Office arranged to have the powerful GBR station 16kc/s controlled by the standard used for MSF.* A photoaphic copy of one of Mr. Pierce's records reproduced in g. C gives results of measurements of both the MSF 60kc/s nd GBR carrier waves. The edge of a dark band represents the rrival time of a particular phase of the wave on a time scale erived from the local standard. The broken line shows the lope for a frequency difference of 1 part in 10^9, and it is clear hat measurements to this accuracy can be obtained in a short ime and that the observations will enable the diurnal change of ropagation time of these long waves to be measured with high recision. Short-time comparisons to this accuracy are of paricular interest in view of the development of atomic standards of requency mentioned by several speakers.

I was interested in Mr. Smith's account of the adoption of Ephemeris Time. We have from the outset expressed the frequencies of MSF in terms of the uniform time system, but it is doubtful whether this has so far given an appreciably greater uniformity because the predicted variation in the period of the earth's rotation was found to be too great. If it can be determined more quickly and predicted with greater accuracy by means of the moon observations we may be able to do better in the future.

The relationship between astronomical time and atomic time will become of great importance, and I do not think we should dismiss the subject with Mr. Smith's comment that an atomic standard provides time interval whereas Ephemeris Time offers a true time system. All time standards provide only time interval, and the basic difference between them is the length of this interval. In case of the earth the interval is so large that it is an easy matter to count the individual completed periods and thus construct a time scale. On the other hand because of the large period the observational accuracy is comparatively low, and in order to define the unit of time interval with a high accuracy it is necessary to take the average of many observations. By averaging over a sufficiently long period almost any order of accuracy can be obtained, but this is of little significance outside the field of astronomy, especially as it is known that the time intervals are not uniform during that period. The large observational errors, as well as the limitations of quartz-crystal clocks, make it difficult to establish a uniform time system and limit the accuracy with which the unit of time can be defined during the comparatively short intervals occupied by most physical experiments to about 1 part in 10^8.

With atomic standards the interval is very small, and in consequence the integration to form a time scale is more difficult; but the precision of setting is correspondingly high, something better

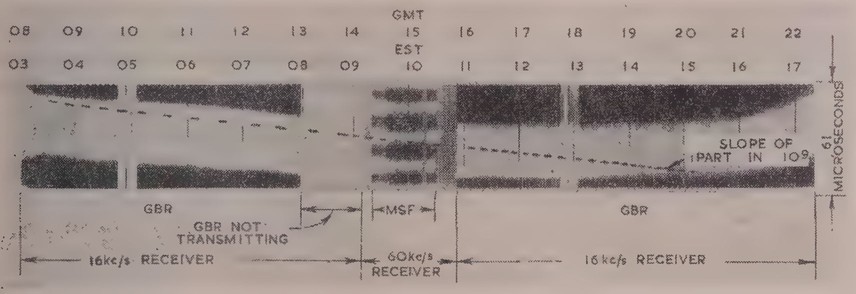


Fig. C.

* Pierce, J. A., Mitchell, H. T., and Essen, L.: "World-wide Frequency and ime Comparisons by means of Radio Transmissions," *Nature*, 1954, 174, p. 922.

than 1 part in 10^8 already having been achieved in the United States. They should therefore enable our working standards—the quartz-crystal clocks—to be calibrated with a precision not previously obtained and moreover in a very short time. The quartz-crystal clocks could then be used with this accuracy for measurements of frequency and time interval.

It is clear from their inherent differences that astronomical standards will always be better for the measurement of long time intervals, while atomic standards promise to be more suitable for short time intervals. It is essential that the relationship between the two be established by appropriate experiments and recognized by international agreement. The independent use of atomic resonances as frequency standards, which appears to have been suggested in some quarters, would lead to much confusion, and should be firmly discouraged.

Some very kind remarks have been made about the quartz-ring oscillators, and I am naturally gratified that they have proved so successful. When I refer to the rings made by the Post Office I am always careful to specify their source, partly in order to acknowledge the work of the Post Office because I know the performance depends to a large extent upon the care taken in manufacture, and partly in order to label them so that the performance can be associated with a particular form of construction. When the rings were first developed the technique of depositing electrodes on quartz was not well established, and I decided to use an independent electrode system. Any movement of the ring within the electrodes produced a large frequency change, and the three-pin support was used to give the necessary geometric location. Now that the electrodes are deposited on the quartz, a more resilient form of mounting is permissible and has some advantages. I can assure Mr. Mitchell that I am not critical of the changes in design which have been made and which have probably resulted in a general improvement of performance, particularly as regards the effect of mechanical shock. At the same time, the drift rate of one of the original two handmade rings used at the N.P.L. was as low as any yet recorded, and two more of these rings reached Australia satisfactorily without any special precautions being taken and are I believe still in use.

I have always felt that the ring is too elaborate for widespread application, and Capt. Booth shares this view; but the intensive efforts that have been made to produce a simpler but equally stable oscillator have not been completely successful. The solution to the problem of the frequency control of transmitters may depend on a wider use of low-frequency standard transmissions and automatic frequency control.

Mr. J. McA. Steele (*in reply*): I am in agreement with Mr. Lea on the desirability of stating the relevant time interval when discussing the limits of error attaching to any frequency comparison. This is generally done in the paper where a precise statement is attempted. In regard to the length of the beat period, the sources of error in the comparison equipment at the N.P.L. are thought to be well known and have negligible effect on the accuracy of comparison, even with beats as short as 2sec.

Owing to the necessarily discrete nature of the counting process in beat-period measurements, slight discrepancies, amounting at the most to 3 parts in 10^{10}, can arise between the traces (*a*), (*b*) and (*c*) of Fig. 4. The relative stability of the ring oscillators, as determined by direct measurement, is usually better than 2 parts in 10^{10} over a period of 24 hours.

In reply to Mr. Griffiths, the standard supplying the frequency generator in Fig. 5 is usually the same as the comparison standard for MSF. Any other standard of comparable stability may, however, be used if it happens to provide a more convenient beat period for manual or automatic measurements.

The phase changes referred to in the Droitwich 200kc/s transmission could only be detected in our measurements by the variation produced in the mean carrier frequency in a short time interval. They are not inconsistent with the frequency differences, amounting at times to 5 parts in 10^9, found between successive comparisons at 4sec intervals with a beat period of length 2sec.

With regard to Mr. Eckersley's suggestion for the control of the mains frequency by reference to a quartz oscillator, it is questionable how far one may regard the frequency as an independent variable. To do so appears to imply that the total generating capacity instantly available at any time greatly exceeds the total demand or potential demand—an operating condition scarcely likely to appeal to the supply engineer.

Mr. H. B. Law (*in reply*): In reply to Dr. Smith-Rose, there have been no interruptions to the MSF transmissions owing to the equipment going out of tolerance or owing to loss of temperature control. The telephone service has a prior call on the long-wave transmitter used for the 60kc/s service; no rebate has been allowed to the National Physical Laboratory for loss of time on the standard frequency service. I would suggest that the N.P.L. gets very good value for money.

Mr. Lea enquires whether the irregularities visible in Fig. 8 are attributable to the oscillations under comparison or to the comparison equipment itself. In the case of the lowest graph, which gives the 24-hour frequency differences between red and blue, comparison errors are negligible; the scatter in this graph is less than 1 part in 10^9. The three remaining graphs involve frequency comparisons with the Post Office standard, each made over a period of some minutes in terms of the 60kc/s transmission; the comparison accuracy is estimated to be rather better than 1 part in 10^9, and the scatter visible in the first and third graphs is consistent with these comparison errors and with the short-period instabilities of the oscillators concerned. The second graph concerns the comparatively unstable white oscillator.

Mr. Griffiths expresses interest in the method used for detecting the presence of the crystal-oven temperature control cycle, but full details cannot be given here. He remarks that he has had very satisfactory results with the "knock-back" binary type of frequency-divider circuit. There are several types of frequency-divider circuit that are good in principle, and my experience suggests that the choice of method is less important than the exercise of proper care in the detailed design work. It so happens that we have made a particularly close study of the action of the multivibrator, and we naturally tend to use it for frequency division.

Capt. Booth remarks that the equipment at Rugby is maintained by the normal transmitting-station staff. I should like to add that the staff concerned have looked after the equipment, which is very different from that ordinarily to be found in a transmitting station, in a most creditable fashion; any success achieved by the MSF service is due in no small measure to their careful work.

317.38 : 621.315.59 : 538.632 : 538.566

Paper No. 1654 M
June 1954

IE APPLICATION OF THE HALL EFFECT IN A SEMI-CONDUCTOR TO THE MEASUREMENT OF POWER IN AN ELECTROMAGNETIC FIELD

By Professor H. E. M. BARLOW, Ph.D., B.Sc.(Eng.), Member.

e paper was first received 17*th November*, 1953, *and in revised form* 26*th January*, 1954. *It was published in June*, 1954, *and was read before the* MEASUREMENTS SECTION 23*rd November*, 1954.)

SUMMARY

he paper deals with the application of the Hall effect in a semi-ductor, like germanium, to the measurement of power transmitted ng a line or a hollow metal waveguide. It is shown that the ression for the instantaneous value of the Hall e.m.f. generated in nall piece of semi-conductor erected in an electromagnetic field is xactly the same form as the expression for the instantaneous power ersing the semi-conductor. Thus the mean value in time of the ll e.m.f. is a direct measure of the power, and this is true for rnating, pulsating or steady fields.

he paper describes various types of wattmeter embodying the fore-ng principle and applied to a wide range of frequencies. The thod is particularly suitable for use when strong magnetic fields are ilable, and heavy-current systems consequently lend themselves to application of this device without the need for current trans-mers. Experiments at 50 c/s and again at 300 Mc/s with an *n*-type manium crystal erected in the annular space between the inner d outer conductors of a coaxial line showed that, for a given power, Hall e.m.f. was the same at the two frequencies, within the accuracy the measurement. Thus there is a prospect of calibrating such an trument at a low frequency and using it, if desired, at a very high quency.

(1) INTRODUCTION

The Hall effect,[1] which was discovered in 1879, occurs when transverse magnetic field is applied to a conductor carrying a ngitudinal current. In such circumstances an e.m.f. is set up the conductor at right angles to the magnetic field and to the rrent. The effect arises from the action of the externally plied field in producing a sideways deflection of the mobile rriers taking part in the conduction process.* Semi-conductors e germanium and silicon show a relatively large Hall effect, and general such materials contain both negative and positive rriers, represented respectively by the so-called excess electrons d holes.[2] The two carriers naturally give oppositely-directed all effects, and to ensure that the resultant Hall e.m.f. is of bstantial magnitude it is important to employ materials in hich conduction is predominantly either by electrons or by oles. Thus germanium of the *n*-type or of the *p*-type is usually tisfactory. The conductivity of a semi-conductor depends pon its impurity content, and when that is reduced the material ecomes a better insulator. On the other hand a conductor hen purified should become a better conductor.

It is the purpose of the paper to describe a new application of he Hall effect in semi-conductors. So far the effect has been mployed only for such purposes as the measurement of steady agnetic field-strengths[3] and for the elucidation of the behaviour f semi-conducting materials.[2]

If a piece of germanium or silicon, for example, is erected in an lectromagnetic field, a Hall e.m.f. will be generated in it, when there is a flow of energy through the field. The orthogonal spatial relationship between the transverse electric- and magnetic-field components provides for the essential conditions required. Thus the mean value, in time, of the Hall e.m.f. produced in that way is a direct measure of the power, and this is true for both steady and alternating fields. Semi-conductors suitable for the purpose have a high resistivity and therefore do not present a serious obstacle to penetration by the field, even at very high frequencies. The real difficulty in applying this device, without ambiguity, to the measurement of power, arises from the rectifier action which a semi-conductor circuit-element tends to introduce, and special precautions are necessary to eliminate so far as possible any residual rectification.

Examples are given of the design of practical instruments for the measurement of power in terms of the Hall e.m.f. generated in a semi-conductor. These instruments range from direct-current and low-frequency types to arrangements suitable for the microwave part of the spectrum. Experiments carried out in relation to this application have established the existence of the Hall effect in *n*-type germanium up to a frequency of 300 Mc/s, and although precise quantitative measurements have still to be made, it is clear from the work already done that the Hall coefficient remains approximately at the same value throughout that frequency band. Thus when a transverse electromagnetic wave is used for the device it seems likely that the arrangement can be calibrated at 50 c/s and used if desired at very high frequencies.

(2) PRINCIPLE OF OPERATION

Suppose that in a piece of semi-conductor, such as germanium or silicon, a magnetic field is set up along the *z*-axis, as defined by a Cartesian co-ordinate system, with current flow at right

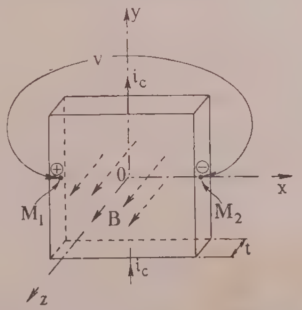


Fig. 1.—Semi-conductor element and co-ordinate system employed.

The positive and negative signs refer to the polarity of the Hall e.m.f. for mobile negative carriers.

angles along the *y*-axis; then an e.m.f. appears, owing to the Hall effect, along the *x*-axis (Fig. 1). This electromotive force at any instant is given by

$$v = \left(\frac{\mathscr{R}}{t}\right) i_c B \text{ volts} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

* The distribution of current over the cross-section of a conductor is actually odified slightly by the magnetic field associated with that current, but the magnitude f this Hall effect is generally so small that it may be ignored without incurring ppreciable error. [See "Electromagnetic Theory," by J. A. Stratton (McGraw-Hill ook Co.), p. 14.]

Dr. Barlow is Pender Professor of Electrical Engineering at University College, ondon.

where B is the instantaneous flux-density in gauss, i_c is the instantaneous current in amperes, $\mathscr{R}$ is the Hall coefficient in volt-centimetres per ampere-gauss, and t is the thickness in centimetres along the z-axis of the sample of material exhibiting the Hall effect.

It can be shown[2] that

$$\mathscr{R} = -\frac{3\pi(n_e b_e^2 - n_h b_h^2)}{8e(n_e b_e + n_h b_h)^2 \times 10^8} \text{ volt-cm/amp-gauss} \quad . \quad (2)$$

where n_e and n_h are respectively the number of excess electrons and holes per cubic centimetre, b_e and b_h are the corresponding mobilities of the excess electrons and holes expressed in square centimetres per volt-second, and e is the electronic charge.

For n-type germanium at room temperature $n_e \gg n_h$, and taking $n_e = 10^{15}$ carriers/cm^3, $b_e = 2\,600$ cm^2/volt-sec, $b_h = 1\,700$ cm^2/volt-sec, with $e = 1{\cdot}6 \times 10^{-19}$ coulomb, we find $\mathscr{R} = -7{\cdot}4 \times 10^{-5}$ volt-cm/amp-gauss.

The power-flow density P at any instant in an electromagnetic field, measured in watts per square centimetre of cross-sectional area of the field, is given by

$$P = EH \text{ watts/cm}^2 \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where E is the instantaneous value of the transverse component of the electric field, in volts per centimetre, and H is the instantaneous value of the transverse component of the magnetic field, in amperes per centimetre.

If, therefore, it is arranged to produce through a piece of semi-conductor, placed in an electromagnetic field, a current i_c arising from the electric field E, then assuming in the first place a purely resistive circuit, the current i_c will be in phase with E and proportional to it. The magnetic flux density B is equal to μH, and if μ, the permeability of the medium, is constant, the Hall electromotive force will be a direct measure of the power-flow density at any given point in the field. It is clear that eqns. (1) and (3) for the Hall electromotive force and the power-flow density, respectively, are, apart from a constant term, exactly equivalent when, for example, i_c corresponds to E, and B to H.

If, for an alternating electromagnetic field, the modulus of i_c is made proportional to the modulus of E and the modulus of B proportional to the modulus of H, with the same phase difference between i_c and E as between B and H, then P is still obtained directly in terms of v. The total power can either be related to the power-flow density over a small area of the field or be obtained by integration.

The orthogonal spatial relationship which normally exists between the transverse components of E and H in an electromagnetic field enables these components to be used to produce the desired effect in the piece of semi-conductor, so that the Hall electromotive force then represents the longitudinal power.

The device is applicable to steady, pulsating or alternating electromagnetic fields. When pulsating or alternating fields are used it is necessary to obtain the true mean value in time of the Hall electromotive force taken over a complete period. Thus, represented as a function of time, the Hall electromotive force itself may be alternating or pulsating according to whether the electromagnetic field consists of, for example, a standing wave or a travelling wave, but in any case its true mean value over a complete period gives a measure of the actual power, taking into account any phase difference between the transverse components of the electric and magnetic fields. Moreover, the direction of the power through the electromagnetic field is given by the sign of the Hall electromotive force and the accuracy of the power measurement is independent of the waveform of the field components. It may be convenient sometimes to operate with a phase difference between B and H or the corresponding current when the field is associated with a metal guide, and to provide for that phase angle by an equal angle of the same sign between i_c and E, when using a pure resistance load for which the field components E and H must be in step (see Fig. 6).

Whilst the device is of general application to the measurement of power or of energy transmitted in an electromagnetic field whether supported by a guide or in free space, it is particularly suitable for use with a transmission line, a cable or a waveguide. In order to obtain a conveniently measurable Hall electromotive force with a power of a few watts, it is desirable when using a semi-conductor like germanium to operate the device in a relatively strong magnetic flux, since the current i_c is limited by the need to avoid any substantial power loss in the semi-conductor and consequent rise in temperature. Moreover, there is a tendency for semi-conductors to saturate in very strong electric fields.[4]

For a free-space plane wave the ratio E/H is 377 ohms, but when conductors are used as a guide it is possible to alter this ratio by increasing H at the expense of E and bringing about a more favourable situation for power measurement by the application of the Hall effect. The magnetic flux density, B, can also be increased relative to H by providing a medium of high permeability over the greater part of the magnetic circuit. There is, however, the possibility of increasing the sensitivity of the device with alternating fields by placing it in a resonant system, such as a resonant window or cavity of the kind employed at ultra-high frequencies.

(3) ELIMINATION OF RECTIFIER ACTION IN SEMI-CONDUCTOR UNIT

For power measurement by the method described the Hall electromotive force set up along the x-axis of the piece of semi-conductor is required, when a current i_c flows through it along the y-axis and a magnetic flux B penetrates it along the z-axis as shown in Fig. 1. The piece of semi-conductor, in so far as its ordinary linear ohmic resistance is concerned, can be conveniently regarded as forming a bridge circuit (Fig. 2) which, when

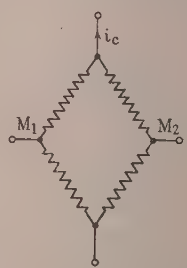


Fig. 2.—Equivalent linear-resistance circuit of semi-conductor element.

electrically balanced, should give no potential difference between the Hall contacts M_1 and M_2 in the absence of any magnetic flux B. It is desirable in the present application to establish this condition as nearly as possible, but it is not essential since any sinusoidal component of the output at M_1,M_2 does not affect the true mean value, and a condenser across these contacts can be used if necessary to by-pass any alternating current arising from an imperfect linear resistance balance. The circulation of such a current will, of course, add to the heating of the piece of semi-conductor and should therefore be avoided as far as possible. There are occasions when it may be worth while to incorporate a filter for this purpose.

Unfortunately, connections to a semi-conductor almost

variably tend to introduce a non-linear resistance,[5] but, even), perfect electrical symmetry of the bridge circuit representing ie semi-conductor unit would be expected to eradicate any ctified current output between the contacts M_1 and M_2. In ich circumstances when the current i_c is alternating it should isplay exact symmetry as between one half-period and another, hilst the corresponding output at M_1 and M_2 should behave milarly. (See Fig. 1.)

Soldered connections to a germanium crystal, particularly if ie solder contains antimony, can be made very largely non-ectifying, or alternatively a welded connection, having the esired properties, may be formed by discharging a condenser hrough a wire in contact with the crystal.[6] It has also been hown that a good so-called non-rectifying connection may be nade to semi-conductors by electrolytically plating their surface t the points required.[7]

Even when every precaution has been taken to establish linear hmic connections to the semi-conductor, it is not always possible o avoid entirely some slight residual rectifying action, but if the oltage applied along the y-axis is not allowed to rise to too igh a value, the rectification effect on the true mean value of he unbalanced e.m.f. appearing between M_1 and M_2, in the ibsence of B, can be made negligible. The need to avoid any ignificant residual rectifier output at the Hall contacts is another reason for increasing the sensitivity of the device as a wattmeter by strengthening B rather than i_c. The electrical balance of the semi-conductor as a bridge circuit can be conveniently carried out, for example in the case of germanium, by attaching the leads to it as symmetrically as possible and then making the final adjustment, whilst current is passing through the crystal, by grinding the edges of the low-resistance arms in the vicinity of the contacts with a piece of carborundum until the nearest possible approach to balance is achieved.

In order to provide for a final adjustment which is easily made should any small change occur in the operating conditions, it is convenient sometimes to use a shunt resistance as shown in Fig. 3(*a*) or a voltage divider as shown in Fig. 3(*b*). These

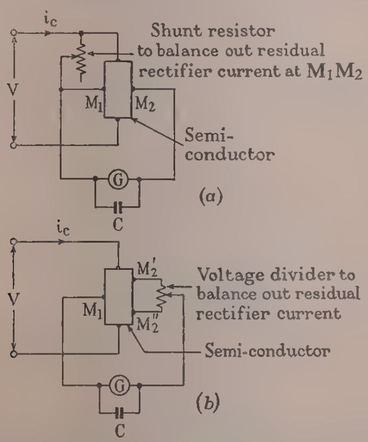


Fig. 3.—Methods of balancing out residual rectifier output at Hall contacts.

controls are placed outside the electromagnetic field, and in the first case the resistance must be very much higher than the internal resistance of the semi-conductor, whereas in the second case the voltage-divider resistance should be low compared with the galvanometer resistance and yet higher than the internal resistance of the semi-conductor between the contacts across which it is connected. Alternatively, an independent rectifier can be used to give an output determined by the voltage at the same point as the semi-conductor unit and matched to compensate any residual rectifier current arising from that unit. There is also the possibility in appropriate circumstances of applying the pole of a permanent magnet or a narrow pencil of light at a particular place on the surface of the semi-conductor unit, so as to bring about an adjustment for electrical symmetry, but this condition is difficult to maintain by that means over a range of voltages applied to the semi-conductor.

When using the device with alternating fields it has been found possible to connect a number of semi-conductor units in cascade so that their individual Hall electromotive forces add together to give an enhanced output for a given power and an increased sensitivity. To do this it is necessary to feed the current i_c to each semi-conductor unit through a capacitance C_1 so that no direct-current circuit is formed by way of the conductor or conductors acting as a guide for the power being measured. This arrangement is shown in Fig. 4. To reduce the effect of C_1 in

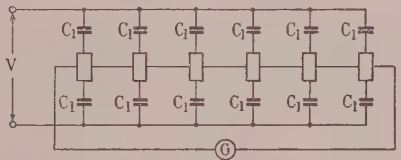


Fig. 4.—Method of connecting a number of semi-conductor elements in cascade, when employed with alternating electromagnetic fields.

changing the phase of i_c relative to the supply voltage V, it is desirable to make the impedance of C_1, at the frequency used, small compared with the resistance of the semi-conductor unit itself. This is easily achieved at high frequencies when the cascade connection of a number of semi-conductor units in the manner described is of particular value. By suitably orienting the individual semi-conductor units in the electromagnetic field, any residual unbalanced rectifier electromotive forces between the Hall contacts can be made to neutralize one another, whilst the required Hall electromotive forces remain additive. Since the true mean value of the Hall electromotive force over a complete period is directly proportional to the power through the semi-conductor, it is clear that the use of a high-resistance moving-coil galvanometer to record that electromotive force yields a linear scale of power. Such a scale is generally preferable to the non-linear scale of, for example, the dynamometer wattmeter, often used at power frequencies.

Fig. 5 shows the Hall e.m.f. as a function of time for an applied alternating voltage and current corresponding to the electric and magnetic fields of an electromagnetic wave when feeding, in one case, a pure-resistance load and, in the other, a pure-reactance load. A small asymmetrical current output, arising from residual rectifier action, is also included on the diagram.

In applying germanium crystals to the measurement of power in the way discussed, it is important to remember that they are sensitive to light and are affected by surface moisture. Any disturbances due to such causes can be substantially eliminated by coating the crystal with an opaque insulating varnish, which is incidentally a water repellent. There is also the problem of heating and the need to avoid any substantial rise of temperature. The number of carriers per cubic centimetre, both positive and negative, varies with temperature, and this affects the value of the Hall coefficient. At ordinary air temperatures the rate of variation is small, but most semi-conductors should not be operated much above room temperatures. At about 70° C the Hall coefficient for germanium begins to fall rapidly, but it

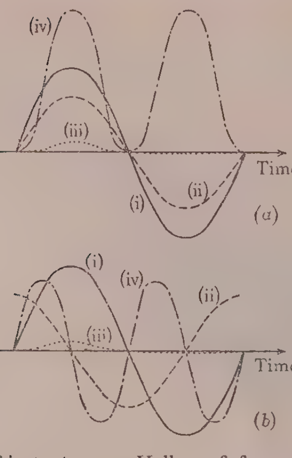


Fig. 5.—Curves of instantaneous Hall e.m.f. for an a.c. circuit feeding (*a*) a pure-resistance load and (*b*) a pure-capacitance load.

(i) Instantaneous load voltage.
(ii) Instantaneous load current.
(iii) Output from rectifier action.
(iv) Output from Hall e.m.f. or power consumed in load.

recovers again on cooling. A small crystal measuring about 1 cm × 0·5 cm × 0·1 cm cannot be allowed to dissipate more than about ¼ watt. Silicon has better properties in this respect.

(4) APPLICATION TO POWER MEASUREMENT

(4.1) Arrangement for Direct Current and Low-Frequency Alternating Current

In the first place the piece of semi-conductor may be erected between the conductors of a parallel-strip transmission line as shown in Fig. 6. The Hall electromotive force is conveniently

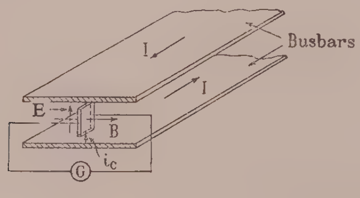


Fig. 6.—Parallel-strip transmission line with semi-conductor unit mounted between the busbars.

measured by a high-resistance galvanometer, by some form of valve voltmeter or, if greater precision is required, by a potentiometer. The arrangement is suitable for either d.c. or a.c. circuits.

The proportionality factor between the mean value of Hall e.m.f. and power, representing the sensitivity of the device, depends upon many things. Using an *n*-type germanium crystal of dimensions 0·5 cm × 0·3 cm × 0·1 cm giving an internal resistance of, say, 650 ohms, supported in air between two parallel-strip conductors each 1 cm wide and separated by a distance of 0·5 cm, a Hall e.m.f. of about 6 microvolts per watt has been obtained with a potential difference across the crystal not exceeding 5 volts.

At power frequencies and within the audio range it is helpful to strengthen the magnetic flux by the use of high-permeability materials, such as laminated iron and its alloys, to form the principal part of the magnetic circuit outside the semi-conductor as shown, for example, in Fig. 7. In this way the sensitivity of the arrangement can be greatly enhanced. Alternatively, an

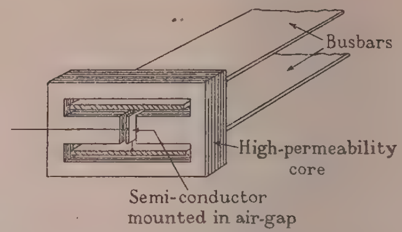


Fig. 7.—Parallel-strip transmission line with semi-conductor unit mounted between the busbars and in the air-gap of a high-permeability core.

arrangement similar to the tong ammeter may be employed so that the semi-conductor is placed in the magnetic field outside instead of between the busbars forming the transmission line, and a voltage circuit for the current i_c is provided by clips attached to opposite points on the busbars as shown in Fig. 8.

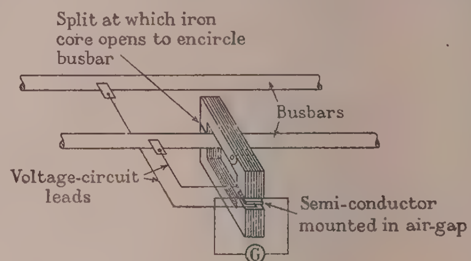


Fig. 8.—Parallel-wire transmission line with high-permeability core encircling one of the conductors and a semi-conductor unit mounted in an air-gap of the magnetic circuit outside the busbars.

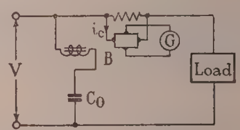


Fig. 9.—Circuit arrangement using a magnetizing coil in parallel with the supply and a series resistor to by-pass current through the semi-conductor unit.

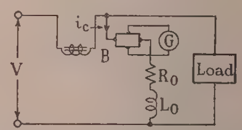


Fig. 10.—Circuit arrangement using a magnetizing coil in series with the supply and a shunt circuit to pass current through the semi-conductor unit.

This arrangement is particularly suitable for heavy-current busbars and avoids the use of a current transformer.

For lower-current circuits at power or audio frequencies the instrument may be set up either as shown in Fig. 9 or as in Fig. 10. In the first case the magnetic flux is provided by a coil with a laminated iron core, connected in series with a condenser C_0 across the power-supply terminals. The magnetic flux is brought into phase with the supply voltage V by adjustment of C_0 whilst the semi-conductor current i_c is made proportional to the main circuit current I. The other arrangement in Fig. 10 interchanges the functions of the shunt and series circuits of the instrument. An inductance L_0 or its equivalent is then generally necessary to

tablish the correct phase relationship between B and i_c. The phasing of the flux and current in the semi-conductor is particularly important when accurate measurements of power at low power factors are required, and the arrangement is conveniently set up by adjustment to give zero mean value of the Hall e.m.f. on a purely capacitive load.

An instrument designed to work in a 230-volt 50-c/s a.c. circuit, using the scheme shown in Fig. 9, with a series resistor of 1 ohm to provide the current through a germanium crystal of dimensions $0{\cdot}4$ cm $\times$ $0{\cdot}2$ cm $\times$ $0{\cdot}1$ cm, gave a sensitivity of about 100 μV/W with a linear scale and an accuracy at least comparable with the corresponding substandard dynamometer instrument. The capacitance of the condenser C_0 was adjusted so that the current through the magnetizing coil was advanced in phase by a small angle relative to the line voltage and sufficiently to bring the flux of the coil into step with that voltage. To avoid errors the laminated iron core must not be allowed to saturate and any residual rectifier output should always be negligible compared with the mean value of the Hall e.m.f. With an inductive-reactance load in which magnetic saturation takes place the current waveform becomes severely peaked, and in such circumstances particular care must be taken to ensure that the maximum voltage applied to the crystal is not sufficient to produce a significant rectified current output at the Hall contacts.

For the power flow along a coaxial line or cable, the arrangement shown in Fig. 11 may be employed and the magnetic flux can then be increased, for a given field strength, by the use of concentric wires or rings of high-permeability materials following the flux path outside the semi-conductor.

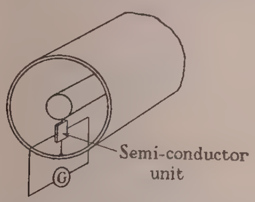


Fig. 11.—Coaxial line with semi-conductor unit mounted between inner and outer conductors.

(4.2) Arrangements for High-Frequency Currents

Coaxial lines are often used at high frequencies, and to make the device (Fig. 11) applicable in such circumstances it is important to give special attention to matching the impedance of the instrument to the characteristic impedance of the line and to ensuring that the leads brought out for measuring the Hall electromotive force are properly disposed in relation to the electromagnetic field. By choosing a semi-conductor of suitable ohmic resistance and by employing the usual reactance matching arrangements such as pistons, tuning screws and diaphragms, impedance discontinuities due to the presence of the semi-conductor, with the leads to it, can be substantially eliminated.

At high frequencies, capacitance connections at opposite sides of the semi-conductor to the inner and outer conductors of the coaxial line are appropriate, and this arrangement enables a number of semi-conductor units to be connected in cascade, adding together their Hall electromotive forces whilst cancelling out residual rectifier effects. The capacitance connection to the semi-conductor also enables phasing adjustments to be made. When a number of semi-conductor units are used in cascade these can conveniently be arranged radially, in a given transverse plane, around the centre conductor of the coaxial line as shown in Fig. 12. The leads to the Hall connections should, so far as possible, always be perpendicular to the electric field. In a coaxial line this is difficult to achieve with the terminal leads to the instrument measuring the Hall electromotive force, and in such cases at high frequencies it is satisfactory to space the leads, where parallel with the electric field, by a quarter of the associated guide wavelength.

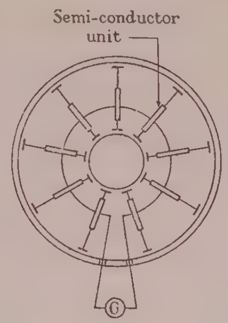


Fig. 12.—Coaxial line with a number of semi-conductor units in cascade and arranged radially around the centre conductor.

Although there are some materials, such as Rhometal, having a relative permeability considerably greater than unity at frequencies as high as 3 000 Mc/s, the use of such materials at these frequencies to increase the magnetic flux penetrating the semi-conductor tends to introduce losses and to present phasing problems. Thus it is usually preferable to employ other means of obtaining the same result.

It is well known that a change in the transverse dimensions of a transmission line brings about a change in the characteristic impedance of the line, and for a given power flow this will enable the magnetic field at certain points to be strengthened at the expense of the electric field. In the case of a coaxial line there is also the fact that the magnetic field is strongest next to the centre conductor. If, therefore, a coaxial line is constricted over a length of half a guide wavelength as shown in Fig. 13, and the semi-conductor placed at the mid-point of this constricted length in close proximity to the centre conductor, greater sensitivity can be achieved without impedance mismatch.

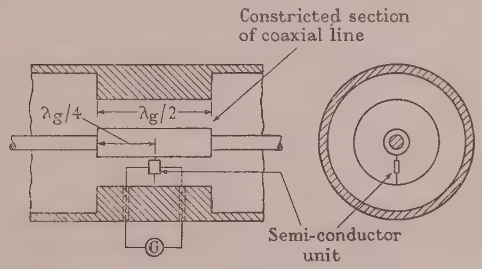


Fig. 13.—Coaxial line with semi-conductor unit mounted at the centre of a constricted section, a half wavelength long.

A line of this kind was constructed to investigate the performance of a germanium crystal applied to the measurement of power at 300 Mc/s, and to compare its behaviour at 50 c/s. For convenience the outer conductor of the coaxial line was made to have the same internal diameter of $1{\cdot}48$ in throughout its length of 6 ft, whilst the inner conductor changed from $0{\cdot}0625$ in outside diameter to $1{\cdot}12$ in outside diameter over a length of $21{\cdot}7$ in. A germanium crystal measuring $0{\cdot}5$ cm $\times$ $0{\cdot}3$ cm $\times$ $0{\cdot}1$ cm was mounted at the centre of the constricted section and arrangements were made to pass power along the line

in either direction whilst maintaining a matched load at the output end. A probe connected to an independent rectifying crystal and galvanometer was also placed at the centre of the constricted length of line so that the electric field strength at that point could be measured. Care was taken to provide complete electric screening of the measuring circuits to prevent spurious disturbances. The rectifier action of the germanium crystal used was not negligible, and the experiments were therefore made with a constant voltage across the crystal whilst observing the change in the deflection of the galvanometer connected to the Hall contacts when the power was interrupted. Thus with a matched load, a power of about 10 watts was first allowed to pass along the line and then stopped by short-circuiting the output end, without change of voltage across the germanium crystal. This condition was established by the use of matching plungers. A direct comparison was in this way obtained between the Hall galvanometer readings for a travelling wave and a standing wave. The measurement cannot be regarded as a precise quantitative one because no facilities were available for evaluating independently the high-frequency power absorbed in the lamp load except by a rough visual assessment of the temperature of the filament, but the observations did show a mean Hall e.m.f. at 300 Mc/s of about 3 μV/W, and for opposite directions of power the galvanometer reading was reversed. Moreover, when the experiment was repeated at 50 c/s much the same result was obtained and the sign of the galvanometer deflection agreed for a given direction of power with the high-frequency measurement.

(4.3) Arrangements for Microwaves

Power flow along the inside of a hollow metal waveguide or along a dielectric guide at ultra-high frequencies can also be measured by applying essentially the same principles. Taking, for example, a metal waveguide of rectangular cross-section supporting an H_{01} wave mode as shown in Fig. 14, it is clear

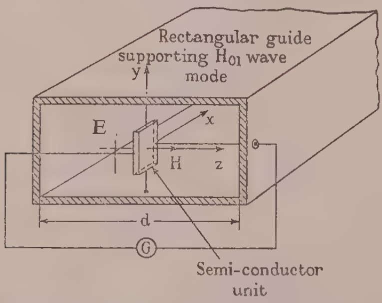


Fig. 14.—Rectangular waveguide supporting an H_{01} wave mode with semi-conductor unit mounted at a point on the axis.

that this arrangement conforms to the basic requirements. The electric field E is entirely transverse and can be made responsible for the current i_c through the semi-conductor, whilst the transverse component of the magnetic field cuts the semi-conductor at right angles. The leads to the Hall contacts are brought out perpendicular with the electric field in the guide, and the usual impedance-matching devices are employed. If required, a stack of semi-conductor units, placed in a given transverse plane near the axis of the guide, can be connected in cascade to build up the Hall e.m.f. for a given power.

With microwaves, advantage can be taken of the cut-off phenomenon to increase the magnetic field relative to the electric field. It is well known that a hollow metal waveguide of rectangular cross-section, whose larger dimension d is less than half a wavelength, ceases to transmit the usual travelling wave and the field in the guide becomes evanescent. The input impedance of a guide operated below cut-off is highly reactive and most of the incident power is reflected at input back to the source. In general, such a guide only transmits power in virtue of a reflected field from the far end of the constricted section, but if both ends are matched by tuning screws and/or diaphragms, so as to cancel out the terminal reactances, then the incident and reflected evanescent fields are built up to permit the whole of the power, apart from losses in the guide, to pass through. In these circumstances the electric field in the evanescent section of the guide is very small and the magnetic field is correspondingly large. Consequently, the conditions in the constricted guide are particularly suitable for the measurement of power passing through it, using the Hall effect in a semi-conductor. The arrangement is shown in Fig. 15, where both

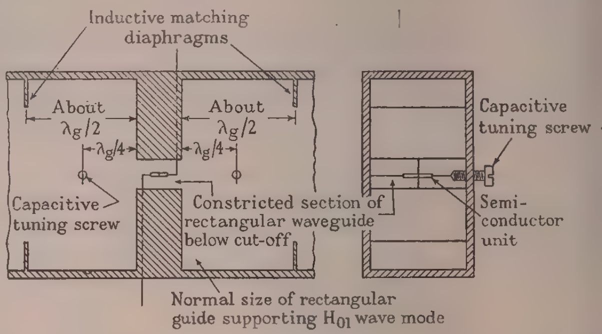


Fig. 15.—Rectangular waveguide supporting an H_{01} wave mode with semi-conductor unit mounted at the centre of a constricted section having impedance matching at input and output.

diaphragms and tuning screws are employed for matching. In this case also a stack of semi-conductor units connected in cascade can be used to increase the sensitivity and nullify residual rectifier effects. The length of the constricted section of guide should preferably only be sufficient to provide satisfactorily for the mounting of the semi-conductor in the evanescent field. If desired, the whole area of the constricted section of guide can be filled with a single piece of semi-conductor; the power flow is then directly calculable from the dimensions and physical constants of the material used.

In the case of n-type germanium having a resistivity of 50 ohm-cm and a permittivity of 16, the skin depth at 3 000 Mc/s is about 0·65 cm.

No experimental information appears to be available at present about the Hall effect at micro-wavelengths, and equipment is now being set up to investigate the behaviour of a germanium crystal for the measurement of power in a hollow metal waveguide according to the plan shown in Fig. 15. At the suggestion of Dr. A. L. Cullen it is proposed independently to attempt to observe the slight elliptical polarization of the electric field which should arise from the Hall effect when a linearly polarized wave is incident on a germanium crystal in the presence of a steady magnetic field along the direction of propagation of the wave.

(5) CONCLUSIONS

The Hall e.m.f. employed to measure power in the way described is necessarily small, and consequently it can be easily disturbed by any residual rectifier action arising in the circuit. This is probably the most difficult single problem that the present development presents. It is not a new one, and already much progress has been made towards a satisfactory solution, but it is probably the most serious limiting factor in the success

a particular application. In general, the best plan is to take ery possible step towards strengthening the magnetic flux and ducing the electric field for a given condition of operation. he final adjustment to eliminate residual rectifier action can en be made by some form of compensating circuit. Other urces of possible inaccuracy do not appear to present serious fficulty under ordinary working conditions.

At low frequencies when high-permeability materials are ailable to strengthen the magnetic flux, the practicability of e method has been established. The application at high fre-uencies has been demonstrated qualitatively, but there is clearly uch work still to be done in the design of a practical instrument n this principle. Should that prove possible there is the pros-ect of a wattmeter which can be calibrated at 50 c/s and used if esired at very high frequencies. This is the more important ecause, at the upper end of the frequency spectrum, power ecomes one of the most significant quantities in electrical easurement and devices at present available for that purpose re by no means ideal.

(6) ACKNOWLEDGMENTS

The author wishes to acknowledge his indebtedness to his olleagues Dr. A. L. Cullen and Mr. H. G. Effemey for many elpful discussions and suggestions concerning this work. He also grateful to Sir Gordon Radley and Dr. E. H. Speight of the Post Office for information on germanium and for the supply of some crystals. The British Thomson Houston Co. and the Standard Telecommunications Laboratories have also been most helpful and co-operative in providing suitable germanium crystals for the experiments.

(7) REFERENCES

(1) Hall, E. H.: "On the New Action of Magnetism on a Permanent Electric Current," *Philosophical Magazine*, 1880, **10**, p. 301.

(2) Shockley, W.: "Electrons and Holes in Semi-Conductors" (D. Van Nostrand and Co., 1950, New York and London).

(3) Pearson, G. L.: "Magnetic Field Strength Meter employing the Hall Effect in Germanium," *Review of Scientific Instruments*, 1948, **19**, p. 263.

(4) Ryder, E. J.: "Mobility of Holes and Electrons in High Electric Fields," *Physical Review*, 1953, **90**, p. 766.

(5) Taylor, W. E., Odell, N. H., and Fan, H. Y.: "Grain Boundary Barriers in Germanium," *ibid.*, 1952, **88**, p. 867.

(6) Mitchell, W. H.: Telecommunications Research Establishment Memorandum, 1952, No. 673.

(7) Pearson, G. L., and Bardeen, J.: "Electrical Properties of Pure Silicon Alloys containing Boron and Phosphorus," *Physical Review*, 1949, **75**, p. 865.

[The discussion on the above paper will be found on page 199.]

621.317.785 : 621.315.59

Paper No. 1778 M
Nov. 195

THE DESIGN OF SEMI-CONDUCTOR WATTMETERS FOR POWER-FREQUENCY AND AUDIO-FREQUENCY APPLICATIONS

By Professor H. E. M. BARLOW, Ph.D., B.Sc.(Eng.), Member.

(*The paper was received* 6*th October*, 1954. *It was published in November*, 1954, *and was read before the* MEASUREMENTS SECTION 23*rd November*, 1954.)

SUMMARY

The wattmeters described may be regarded as typical instruments, and depend for their operation upon measurements of the Hall e.m.f. produced by a germanium crystal when placed in a magnetic field. The crystal carries a definite fraction of the load current and the magnetic field is made proportional to the supply voltage. The power-frequency instrument for use up to 150c/s incorporates an iron-core magnetizing coil, while the audio-frequency design, capable of operation up to 20kc/s, employs only air cores and includes careful screening arrangements. Experience has shown that high-permeability materials must be avoided for precision measurements of power by this means, except over a comparatively narrow frequency band. The audio-frequency instrument, which gives about 2cm deflection for 50mW, is shown to read 3·4% low at 15kc/s and 5·3% low at 20kc/s.

LIST OF PRINCIPAL SYMBOLS

v = Hall e.m.f.
$\mathscr{R}$ = Hall coefficient.
i_c = Current through semi-conductor element.
B = Flux density applied to the semi-conductor element.
t = Thickness of semi-conductor element.
x, y, z = Cartesian co-ordinate axes.
L_0R_0 = Inductance and resistance respectively of magnetizing coil.
R_s = Series resistance in load circuit.
L_bR_b = Inductance and resistance respectively of compensating coil.
R_a = Swamping resistance in series with magnetizing coil.
R_m = Resistance for adjusting electrical balance of crystal.
V = Supply voltage.
I = Load current.
R_f = Feedback resistance in galvanometer amplifier circuit.
ϕ = Phase angle of load.
θ = Angle of lag of flux behind supply voltage.
ψ = Angle of lag of semi-conductor current behind load current.
A = Constant.
$\alpha = (\theta - \psi)$.
P = Mean power.
n_0 and n_c = Percentage errors.

(1) INTRODUCTION

In a recent paper[1] the application of the Hall effect in a semi-conductor to the measurement of power in an electromagnetic field was described in principle. Instruments based on this idea have since been designed for use at frequencies up to 20kc/s, and details of this development are now given, together with typical performance characteristics.

Dr. Barlow is Pender Professor of Electrical Engineering at University College, London.

(2) BASIC PRINCIPLES AND GENERAL CONSIDERATION AFFECTING THE APPLICATION

A conductor, or preferably a semi-conductor element a shown in Fig. 1, erected in an electromagnetic field, produces Hall e.m.f. at any instant along the x-axis given by

$$v = \left(\frac{\mathscr{R}}{t}\right) i_c B \text{ volts}$$

If the current i_c through the element and the applied flu density B are made instantaneously proportional, one to th

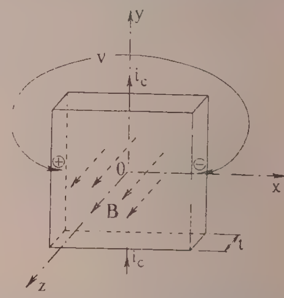


Fig. 1.—Semi-conductor element.

The positive and negative signs refer to the polarity of the Hall e.m.f. for mobi negative carriers.

electric and the other to the magnetic component of the electro magnetic field, the power will be proportional to the instantaneou Hall e.m.f.

Attention has previously been drawn to the desirability c operating the semi-conductor element in this application wit the strongest possible magnetic field for a given power transmitted this helps to ensure that residual rectifier effects arising from th element are relatively negligible.

At frequencies up to and including the audio band it is cor venient to use coils as a means of strengthening the magneti field. Shunt and series circuits through the instrument ar employed much in the same way as in the conventional types c wattmeter, and the problems of design are very similar, fc example, to those of the dynamometer instrument. For lo frequencies (below about 150c/s) high-permeability material can sometimes with advantage be introduced to reduce th reluctance of the magnetic circuit of the magnetizing coil, bu at audio frequencies the waveform distortion that inevitabl arises from the use of such materials prohibits their applicatio in this way. Indeed, at the lowest frequencies great care mus be taken to avoid any approach to saturation of the core materials since flux-wave distortion can be a source of serious erro particularly when dealing with low-power-factor loads.

The magnetizing-coil circuit should preferably carry a shur current, because the conditions in this circuit are then determine by the design of the instrument and are independent of the loac With a series magnetizing coil any flux distortion can be greatl

nagnified by a capacitive load. The basic circuit of the instrument is shown in Fig. 2. The shunt magnetizing coil L_0R_0 producing the flux B in which the semi-conductor element is

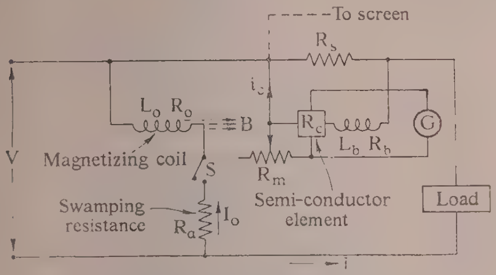


Fig. 2.—Circuit of the semi-conductor wattmeter.
The screen is used only in the a.f. instrument.
L_0R_0 = Magnetizing coil.
R_a = Swamping resistor.

immersed has a large swamping resistor R_a in series with it, so that at the highest frequency employed the impedance of this circuit does not increase significantly. A series resistor R_s in the load circuit provides a voltage to drive a small current i_c through the semi-conductor element by way of the compensating inductance L_bR_b. This compensating inductance is employed to bring i_c into phase with B on a pure resistive load.

The semi-conductor element is an n-type germanium crystal of approximate dimensions $0{\cdot}3\,\text{cm} \times 0{\cdot}25\,\text{cm} \times 0{\cdot}03\,\text{cm}$, to which four wires are soldered forming a balanced-bridge network within the crystal. Semi-conductors in a circuit tend to introduce rectification, and it is most important that any effects of this kind should be eliminated so far as possible, because they produce an output at the Hall contacts in the absence of any externally applied magnetic field. The crystal as a bridge circuit should therefore be balanced to alternating as well as to direct currents. The crystal is best made up by soldering wires to it as symmetrically as possible and then, while alternating and direct currents are alternately passed through it, filing the soldered connections with a piece of carborundum until the nearest approach to electrical balance is achieved. When the crystal is in operation for power measurement the final adjustment can then be made by a variable resistance R_m across one arm of the crystal bridge circuit. It is desirable to set this resistance for zero output from the Hall contacts of the crystal when the current through the magnetizing coil is temporarily interrupted.

(3) CRITERIA GOVERNING DESIGN

It is necessary in the first place to ensure that the shunt current and the voltage drop in the instrument do not seriously disturb the conditions in the circuit under examination. This becomes a more difficult requirement to meet as the level of power to be measured decreases. Secondly, the shunt current must not change significantly in amplitude for a given applied voltage over the whole of the frequency range for which operation is required. Neglecting stray capacitances, it is easy to show that the percentage increase in the modulus of the impedance of the magnetizing-coil circuit over its zero-frequency value is given by

$$50\left(\frac{\omega L_0}{R_0 + R_a}\right)^2 \text{ per cent}$$

and if, at the highest frequency of operation, $\hat{\omega}$, the increase in impedance is limited to 1%, we have

$$\left(\frac{R_0 + R_a}{\hat{\omega} L_0}\right) = 7{\cdot}07 \quad . \quad . \quad . \quad . \quad . \quad (1)$$

At the upper end of the audio-frequency scale stray capacitances may become important and screening arrangements are necessary. The change in phase of the magnetizing current and of the flux produced by it can readily be compensated by providing a circuit for the crystal current i_c of similar impedance characteristic. Thus, on the assumption that $i_c \ll I$, the resistance/inductance ratio of the crystal circuit must be the same as that of the magnetizing coil circuit, and we have

$$\frac{R_c + R_b}{L_b} = \frac{R_a + R_0}{L_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The inclusion of L_bR_b in the crystal circuit for the purpose of phasing i_c correctly adds to the impedance of this circuit, and therefore only relatively small adjustments by this means are permissible. A 1% change in impedance of the magnetizing-coil circuit corresponds to a 1% change in the impedance of the crystal circuit when correct phasing has been achieved.

It is generally desirable to use small crystals, because this reduces, not only the requisite core area of the magnetizing coil and consequently its inductance, but also the internal resistance R_c of the crystal itself, enabling sufficient current i_c to flow with a relatively small resistance in the load circuit.

(4) MEASUREMENT OF THE HALL E.M.F.

The true mean value of the Hall e.m.f., as a measure of the power in the load circuit, is necessarily small in most applications. Over the 50c/s frequency range the use of a magnetizing coil with a high-permeability core enables an ordinary moving-coil instrument to be employed to record the Hall e.m.f., but in the audio band, when air-core magnetizing coils are necessary, a galvanometer amplifier of some kind must be introduced. In such circumstances the sensitivity of the device is generally of the order of a few microvolts per watt, and a very convenient instrument for measuring these small Hall e.m.f.'s is that devised by Hill,[2] incorporating a differentially operated photo-voltaic cell with a series feedback resistance R_f, as shown in Fig. 3. The

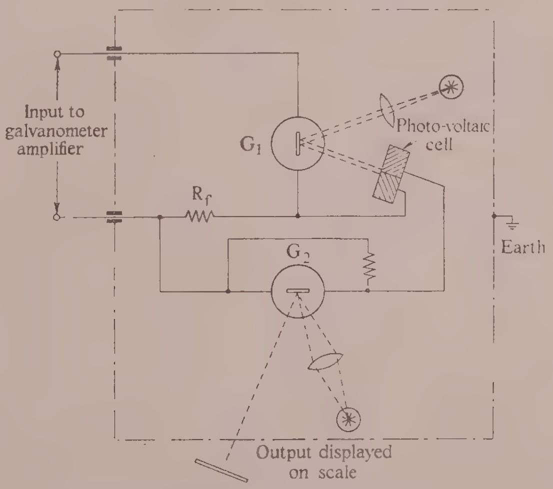


Fig. 3.—Photocell-amplifier unit for measurement of Hall e.m.f.
R_f = Feedback resistor.
—·— Metal screen.

resistance of the crystal wherein the Hall e.m.f. is generated is of the order of 1000 ohms, and the input resistance of the instrument measuring this e.m.f. should be considerably higher than that figure if the voltage drop in the crystal is to be relatively insignificant. At the same time, with such resistances the measuring instrument will be subject to heavy electrical damping,

and a compromise must therefore be made. The conditions are, in fact, very similar to those of the thermocouple instrument for the measurement of current.

When the selenium-photocell galvanometer amplifier (Fig. 3) is being set up arrangements are made for light reflected from the mirror of the first galvanometer, G_1, to be incident on the cell so that an output is obtained only when the light beam is displaced from its mid-position. In those circumstances the second galvanometer, G_2, is deflected while feeding back current through R_f, thereby increasing the effective impedance at the input terminals. To avoid disturbance due to mechanical vibrations the galvanometer G_1 should be of the balanced type developed by Downing.[3] When the resistance of each galvanometer was 450 ohms and R_f was 95 ohms, the input impedance of the apparatus was found to be 34 kilohms, and its calibration curve is shown in Fig. 4.

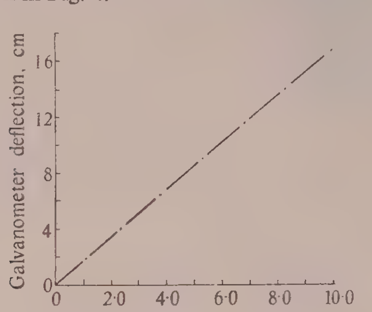


Fig. 4.—Calibration curve of galvanometer amplifier.
Effective input resistance = 34 kilohms.
Resistance of G_1 and G_2 = 450 ohms.
Shunt across G_2 = 10 kilohms.
R_f = 95 ohms.

(5) PHASE-ANGLE ERRORS

Suppose that the supply voltage V is given by $V = V_0 \sin \omega t$ and the load-circuit current by $I = I_0 \sin (\omega t + \phi)$.

The corresponding flux density produced by the magnetizing coil can be expressed as

$$B = B_0 \sin (\omega t - \theta) = A_1 V_0 \sin (\omega t - \theta)$$

and the current i_c through the crystal as

$$i_c = I_c \sin (\omega t + \phi - \psi) = A_2 I_0 \sin (\omega t + \phi - \psi)$$

We then have the true mean value of the Hall e.m.f. as

$$v = \text{Mean}\,(A_3 B i_c) = \frac{A V_0 I_0}{2}[\cos \phi \cos (\theta - \psi) - \sin \phi \sin (\theta - \psi)]$$

or

$$v = \frac{A V_0 I_0}{2} \cos (\phi + \alpha) \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The power in the load circuit, neglecting the loss in R_s is

$$P = \text{Mean}\,(VI) = \frac{V_0 I_0}{2} \cos \phi \quad . \quad . \quad . \quad . \quad (4)$$

From eqns. (3) and (4) it is clear that when $\alpha = 0$, i.e. $\psi = \theta$, then $v = AP$, so that the Hall e.m.f. is directly proportional to the power. For a finite positive value of α with $\phi = 0$, i.e. a purely resistive load, the value of v is slightly smaller than it should be and the instrument reads low; with $\phi = +90°$, i.e. a purely capacitive load, the instrument reads negative instead of zero, and with $\phi = -90°$, i.e. a purely inductive load, the instrument reads positive instead of zero.

In general $\quad \frac{v}{P} = A[\cos \alpha - \tan \phi \sin \alpha]$

from which we find

$$\frac{d}{d\alpha}\left(\frac{v}{P}\right) = A[-\sin \alpha - \cos \alpha \tan \phi]$$

Since α is in any case very small, we get

$$\frac{d}{d\alpha}\left(\frac{v}{P}\right) \simeq -A[\alpha + \tan \phi] \quad . \quad . \quad . \quad . \quad (5)$$

For a resistive load $\phi \simeq 0$ and $d/d\alpha(v/P) \simeq -A\alpha$, which is small. Thus, in this case small phase-angle errors do not seriously affect the accuracy of the instrument.

For a capacitive or inductive load $\phi \simeq \pm 90°$ and $d/d\alpha(v/P) \simeq -A \tan \phi$. Comparatively large errors may therefore arise with loads of low power factor when $\tan \phi$ is large. The adjustment for correct phasing is conveniently made on a purely capacitive load.

High-permeability materials in the core of the magnetizing coil tend to increase θ, as does the self-capacitance of the winding. These factors must be taken into account when adjusting L_b to make $\psi = \theta$.

(6) PERCENTAGE ERROR IN POWER MEASUREMENT DUE TO INCREASE OF IMPEDANCE WITH FREQUENCY OF MAGNETIZING COIL AND OF CRYSTAL CIRCUITS

Suppose that at the highest frequency $\hat{f}$ for which the instrument is designed the current I_0 falls by n_0% and the current i_c falls by n_c%, owing in each case to the increased impedance of the circuit concerned over its d.c. value. The reading of the instrument at $\hat{f}$ will be proportional to

$$\left[I_0\left(1 - \frac{n_0}{100}\right)\right]\left[i_c\left(1 - \frac{n_c}{100}\right)\right] \simeq I_0 i_c\left[1 - \left(\frac{n_0 + n_c}{100}\right)\right]$$

neglecting the second-order term.

Hence the percentage error in the power measurement arising in this way is $-(n_0 + n_c)$. Thus, if $n_c = n_0 = 1\%$, we should expect the instrument to read 2% low at the highest frequency. An allowance can be made in the usual way for the power consumed by the instrument itself.

(7) DETAILS OF DESIGN OF A WATTMETER FOR USE UP TO 150 c/s

The instrument was designed for application in a 230-volt circuit. The magnetizing coil with its core of Stalloy stampings is shown in Fig. 5(*a*). The 2-section winding connected in series gave $L_0 = 0{\cdot}7$H and $R_0 = 317$ ohm, and had an air-gap about 0·05 cm long, housing the germanium crystal. A swamping resistance, R_a, of 13·5 kilohms was included in the magnetizing-coil circuit, giving the upper value of the current $I_0 \simeq V/(R_0 + R_a) = 16{\cdot}66$mA, and an increase of impedance at $\hat{f} = 150$c/s of $50[\hat{\omega} L_0/(R_0 + R_a)]^2 = 0{\cdot}12\%$.

The crystal resistance, R_c, was 1 030 ohms, and proper phase compensation was found to be obtained experimentally with $L_b = 51{\cdot}5$mH and $R_b = 4{\cdot}5$ohms, using a closed Mumetal-core choking coil as shown in Fig. 5(*b*). Eqn. (2) yields $L_b = 52{\cdot}3$mH. For adjustment of electrical balance of the crystal, R_m = 32 kilohms.

The voltage across a crystal of the size employed in this application should not exceed about one or two volts, and R_s is chosen accordingly for the particular upper limit of load

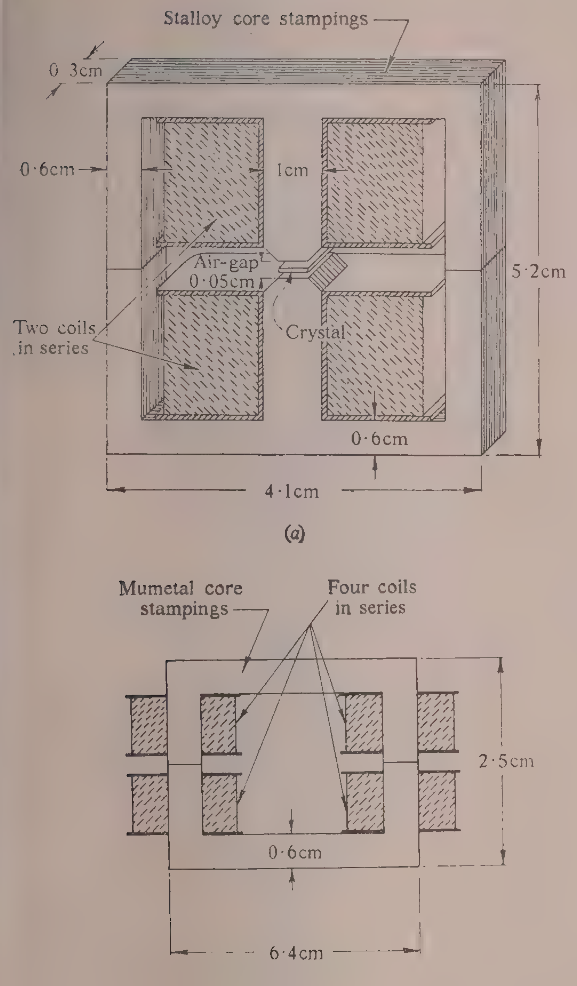


Fig. 5.—Arrangements for use up to 150 c/s.

(a) Magnetic circuit and crystal; each coil has 3 000 turns of 41 S.W.G. wire, giving $L_0 = 0{\cdot}7$H and $R_0 = 317$ ohms.
(b) Compensating inductance; each coil has 320 turns of 24 S.W.G. wire, giving $L_b = 51$ mH and $R_b = 4{\cdot}5$ ohms.

current. Thus, for load currents up to 1 amp, if R_s is 1 ohm the maximum power dissipation in the crystal is rather less than 1 mW. At the same time, the power consumed by the magnetizing-coil circuit is about 3·8 watts, so that with a purely resistive load taking 1 amp the power consumed in the instrument is 1·65%, a figure that compares not unfavourably with many dynamometer instruments.

In recording the true mean value of the Hall e.m.f. as a measure of power, it is convenient to use a unipivot moving-coil instrument. Such an instrument, having a resistance of 879 ohms, gave indications of power as shown in the calibration curve Fig. 6(a). Allowing for the voltage drop in the internal resistance of the crystal, Hall e.m.f.'s of 90 μV/watt were obtained with this arrangement.

At 100 c/s the instrument was found to read 3·4% low and at the highest frequency of 150 c/s the error rose to 7% [see Fig. 6(b)]. These figures are very much higher than would be expected from calculations based on the considerations already discussed, and the reason for this divergence was found to be the combined effect of the high-permeability cores of the magnetizing coil and of the balancing inductance in distorting the flux waveform. The instrument described in the next Section illustrates what can be achieved using air-core coils, and the results emphasize the importance of adopting this arrangement when high precision is required over a wide range of frequencies. In any case, the effective increase in Hall e.m.f. per watt as a result of using the high-permeability cores is only about 2 : 1.

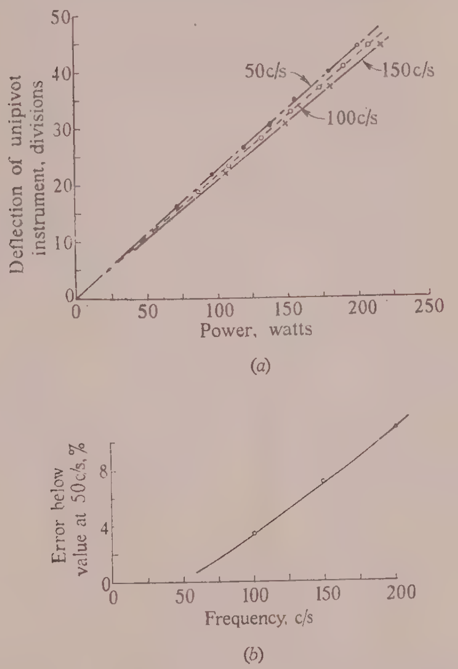


Fig. 6.—Characteristics of power-frequency wattmeter.

(a) Calibration.
(b) Frequency/error.

(8) DETAILS OF DESIGN OF AN A.F. WATTMETER FOR USE UP TO 20 KC/S

At audio frequencies up to 20 kc/s two important additional considerations arise, namely the need to use an air-core magnetizing coil of low self-capacitance and the provision of effective screening. The design of the instrument was based on its application to the measurement of about 1 watt in a circuit having not more than 20 volts across the terminals. The air-core magnetizing coil was represented by a winding with 10 sections in series giving a total L_0 of 29·8 mH and R_0 of 665 ohms, the crystal being mounted at the centre of the core, as shown in Fig. 7(a). This unit was enclosed in a metal box [see Fig. 7(b)] which was itself housed, with the other circuit components, inside another larger screening box, as shown in Fig. 7(c). Non-inductive wire-wound resistors were used throughout, and a switch S with a thermocouple for current measurement were included in the magnetizing-coil circuit. The swamping resistance R_a was 21 kilohms, so that the increase in impedance of this circuit at 20 kc/s amounted to 1·5% and the current did not exceed about 0·92 mA.

The crystal resistance R_c was 1 100 ohms, and phase compensation was obtained with $L_b = 1{\cdot}52$ mH and $R_b = 14$ ohms—values which agree very closely with calculation. The compensating inductance was a small air-core coil included in the inner screen box with the crystal unit. The performance of the

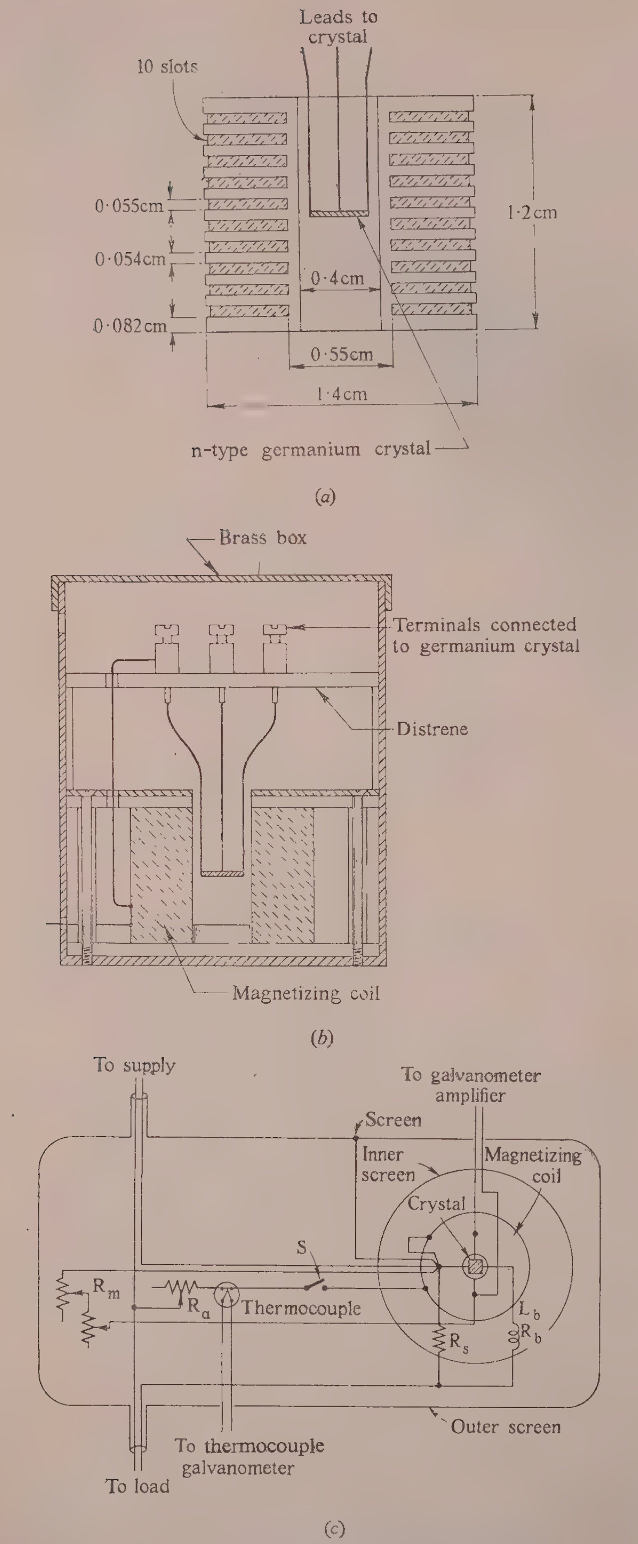


Fig. 7.—An a.f. wattmeter for use up to 20kc/s:
(a) Magnetizing coil; each slot contains 300 turns of No. 47S.W.G. wire; the crystal is of n-type germanium, and R_c = 1 100 ohms.
(b) Magnetizing-coil and crystal assembly; L_0 = 29·8mH and R_0 = 665 ohms.
(c) Screen assembly and components.

device was examined using either a non-inductive load resistance of 590 ohms or a 0·02μF capacitor, both of which were completely screened.

Thus, with a voltage of 20 volts applied to the purely resistive-load circuit, I was 33·9mA and the power consumed was about 0·68 watt. With R_s = 9·92 ohms the voltage drop in the instrument did not exceed 0·34 volt, or 1·7% of the supply voltage, and the shunt magnetizing current was 2·7% of the load current. A photocell galvanometer amplifier (Fig. 3) was used to measure the Hall e.m.f., and the calibration curve for the equipment as a whole is shown in Fig. 8(a), with percentage errors over the frequency range up to 20kc/s in Fig. 8(b). These errors are higher than the calculated values, but are of the same order.

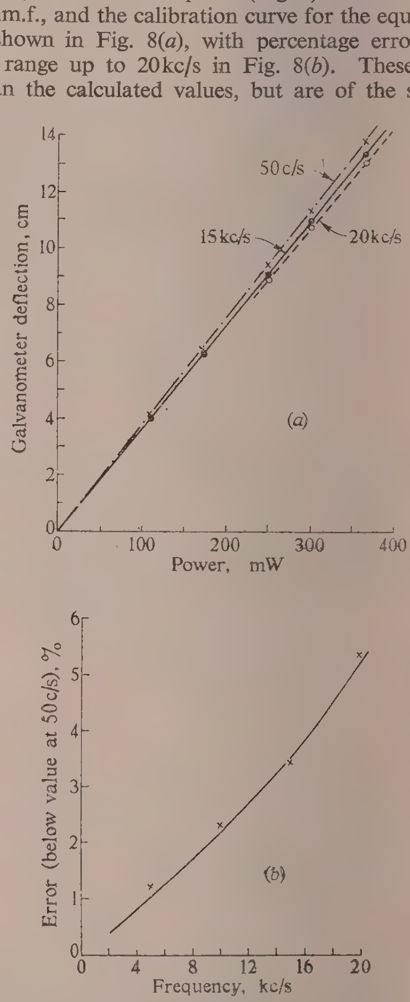


Fig. 8.—Characteristics of a.f. wattmeter.
(a) Calibration.
(b) Frequency/error.

To reduce the effect of stray capacitances in the circuit of the instrument it was found necessary to connect to the screen one side of the input to the crystal. This ensured that no point on the local crystal circuit, galvanometer amplifier or magnetizing coil was far removed from the screen potential. It must be remembered that there is a pronounced second harmonic of the supply frequency in the output from the crystal when power is being measured, and while capacitance between the two conductors connected to the Hall contacts is of no great consequence, stray capacitance between each of these conductors and the screen may be important, particularly if it produces unsymmetrical loading of the crystal. The resistance of the crystal itself was 1100 ohms when R_m was 32 kilohms, and since a

pacitance of 0·0005 μF has an impedance of 8 kilohms at)kc/s (twice the upper limit of the supply frequency), it is early important to reduce to a minimum the capacitance :tween the screen and the galvanometer amplifier leads and to ılance the residual capacitances of this kind so far as possible.

The operating currents of the galvanometer amplifier are :cessarily small, and a good standard of insulation is therefore quired to avoid leakage. Coating the various parts of the)paratus with shellac was helpful in this respect.

When a measurement was being made with this semi-conductor attmeter it was found desirable, before taking a reading, to iterrupt the magnetizing current with the switch S and to ıeck that any rectifier effect of the crystal was properly balanced ut by adjustment of R_m. Alternatively, the change in the alvanometer-amplifier reading when the switch S was closed ould be used as a measure of the power.

(9) CONCLUSIONS

The wattmeter instruments described in the paper offer certain dvantages over other types, while at the same time they suffer om some features that compare unfavourably. The small Iall e.m.f.'s generated when it is required to measure a few atts of power is a distinct handicap, but for heavy-current ystems where kilowatts are involved the problem does not arise. 'hus the power output of a large star-connected 3-phase transormer could readily be measured by bringing out the three eutral conductors before linking them together, and using one o generate the magnetic field required for the crystal while a mall shunt current was passed through the crystal itself. A everse power relay could also be constructed on the same rinciple. As wattmeters, the instruments employing air-core magnetizing coils with phase compensation are capable of a high degree of accuracy at any power factor and any frequency within the range for which they are designed. The power can be recorded remotely, if desired, because the measurement of the Hall e.m.f. is quite independent of the rest of the instrument. The wattmeter also has the advantage of giving a linear scale of power. The dissipation in the crystal is so small that no difficulty has arisen from heating effects, and no significant ageing of crystals has been observed. Disturbances due to eddy currents induced in the crystal also appear to be negligible.

(10) ACKNOWLEDGMENTS

The author desires to record his thanks to the R. W. Paul Instrument Fund for the financial assistance given in this development, and to acknowledge his indebtedness to his colleagues, Prof. F. Brailsford, Dr. A. L. Cullen, Mr. H. G. Effemey and Mr. L. D. Tebbenham for help and advice in various ways. He is also grateful to the British Scientific Instrument Research Association for information on the subject of d.c. amplifiers applied to measurements.

(11) REFERENCES

(1) Barlow, H. E. M.: "The Application of the Hall Effect in a Semi-Conductor to the Measurement of Power in an Electromagnetic Field" (see page 179).

(2) Hill, A. V.: "Moving-Coil Galvanometers of Short Period and their Amplification," *Journal of Scientific Instruments*, 1948, **25**, p. 225.

(3) Downing, A. C.: "Moving-Coil Galvanometer of Extreme Sensitivity," *ibid.*, 1953, **30**, p. 40.

[The discussion on the above paper will be found on page 199.]

621.317.382.029.4 : 621.317.784.082.742

Paper No. 1653 M
Apr. 1954

AUDIO-FREQUENCY POWER MEASUREMENTS BY DYNAMOMETER WATTMETERS

By A. H. M. ARNOLD, Ph.D., D.Eng., Associate Member.

(*The paper was first received* 23*rd November*, 1953, *and in revised form* 25*th January*, 1954. *It was published in April*, 1954, *and was read before the* MEASUREMENTS SECTION 23*rd November*, 1954.)

SUMMARY

Dynamometer wattmeters can be designed and constructed for the measurement of power at frequencies up to about 20 kc/s. The achievement of an accuracy at the higher audio frequencies comparable to that which is possible at power frequencies is dependent first upon certain precautions being taken with regard to the screening of the load and the wattmeter, and secondly upon the possibility of an accurate determination of the frequency-dependent errors of the wattmeter.

The screening necessary is described, and an account is given of the methods adopted at the N.P.L. to calibrate wattmeters.

For any dynamometer a frequency is eventually reached at which a significant instrument deflection is obtained with voltage only or current only applied to the terminals. This is considered to be the useful upper limit of frequency for that instrument.

LIST OF PRINCIPAL SYMBOLS

T = Electromagnetic torque of dynamometer.
θ = Deflection of wattmeter index.
P_W = Power indicated by wattmeter.
P_L = Power consumed in load.
f = Frequency, c/s.
V = Voltage between load terminals.
V_e = Voltage between load screen and the load line not containing the wattmeter current-coil.
I_f = Fixed-coil current.
I_m = Moving-coil current.
I_e = Current from load screen or inner load screen to earth.
I = Load current.
R'_f = Fixed-coil resistance on direct current.
R_f = Fixed-coil resistance on alternating current.
R'_m = Moving-coil resistance on direct current.
R_m = Moving-coil resistance on alternating current.
R_s = Series resistance of wattmeter voltage circuit.
$r = R'_m + R_s$
R = Load resistance.
R_A and R_B = Two parts of load resistance.
X_A and X_B = Two parts of load reactance.
L_f = Fixed-coil inductance on alternating current.
L_m = Moving-coil inductance on alternating current.
M = Mutual inductance between fixed and moving coils.
M_q = Mutual inductance used to provide voltage between quadrants in zero-power-factor test.
C_f = Self-capacitance of fixed coils.
C_m = Self-capacitance of moving coil.
C_s = Self-capacitance of series resistor, including any added capacitance.
C_e = Effective capacitance to earth from mid-point of series resistor, including any added capacitance.
Φ = Magnetic flux produced by I_f.
ϕ = Phase angle between V and I.

Various coefficients are defined as they are introduced. Currents and voltages are generally assumed to vary sinusoidally with time and the dot notation is employed, thus $V . I = VI \cos \phi$.

(1) INTRODUCTION

The use of portable dynamometer wattmeters for measuring power on direct or alternating current at frequencies from about 20 to 100 c/s, usually known as power frequencies, is well established. An accuracy of 1% of the indication may be achieved without difficulty over a large part of the range of a good instrument, the only correction necessary being for the power dissipated in the fixed coils or in the moving coil and series resistor, according to the method of connecting the wattmeter into the circuit. A somewhat better accuracy may be obtained if corrections are made for instrument errors, while still better accuracy is possible with sensitive instruments of a semi-portable type.

The instrument error on direct current may be determined with the aid of a potentiometer and auxiliary equipment. Additional tests on alternating current must then be made to determine the d.c./a.c. sensitivity ratio and the power-factor error. An alternative procedure is adopted at the N.P.L. of carrying out the full tests on alternating current and, if required, the d.c./a.c. sensitivity is determined by means of an additional d.c. test. The a.c. tests are usually made by comparing the indication of the test instrument with that of a standard wattmeter, using the "artificial load" method, although this is sometimes supplemented by real-load tests. At power frequencies the corrections determined from an artificial-load test are applicable when the wattmeter is used for real-load measurements. This may not always be the case at audio frequencies.

The sensitivity of the dynamometer is not decreased when the frequency is increased, although the impedance of the current circuit is increased, as also are the instrument errors—especially the power-factor error. There is an increasing use of these instruments in the audio-frequency range, and it has to be considered whether by a proper design, or by a careful determination of errors, it is possible to achieve accuracy comparable to that at power frequencies. No change in measurement technique is usually necessary up to a frequency of 1 kc/s, but in the upper audio-frequency range errors from stray-capacitance currents may become significant, and screening of the wattmeter and the load may become necessary. A correction may be applied for the power-factor error or compensating circuits may be introduced to reduce it to a negligible value. In either case, the effect is to bring into prominence other frequency-dependent errors, which are normally small compared to the power-factor error.

The electrical errors of a wattmeter may be analysed into four types as follows:

Eqn. (1) gives the electromagnetic torque, T, of a dynamometer for a current I_m in the moving coil and a magnetic flux Φ produced by the current I_f in the fixed coils.

$$T = K\Phi . I_m \frac{dM}{d\theta} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The paper is an official communication from the National Physical Laboratory.

K is a constant and $dM/d\theta$ is the differential coefficient of the mutual inductance, M, between the fixed and moving coils with respect to deflection, θ. Φ is primarily dependent on the load current, I, but may also be dependent to a small extent on the load voltage, V. I_m is primarily dependent on the load voltage, but may also be dependent to a small extent on the load current.

We therefore write

$$\Phi = I(A + jB) + V(a + jb) \quad . \quad . \quad . \quad . \quad (2)$$

$$I_m = V(C + jD) + I(c + jd) \quad . \quad . \quad . \quad . \quad (3)$$

where the coefficients A, a, etc., are all frequency-dependent.

Eqn. (4) may then be written down for $\Phi \,.\, I_m$ where ϕ is the angle between the vectors V and I, and P_L or $VI\cos\phi$ is the load power.

$$\begin{aligned}\Phi \,.\, I_m &= (AC + BD + ac + bd)P_L \\ &\quad + (BC - AD - bc + ad)VI\sin\phi \\ &\quad + (aC + bD)V^2 + (Ac + Bd)I^2 \quad . \quad . \quad (4)\end{aligned}$$

In addition to the electromagnetic torque, there may be a small electrostatic torque which may be similarly resolved. Eqn. (5) may be written down for the wattmeter indication, P_W, where the coefficients combine the effects of both electromagnetic and electrostatic torques

$$P_W = k_1P_L + k_2\, VI\sin\phi + k_3V^2 + k_4I^2 \quad . \quad (5)$$

The proper calibration of a wattmeter consists of the determination of the coefficients k_1, k_2, k_3 and k_4 for all frequencies, deflections and other conditions which are relevant.

The coefficient k_1 depends on the error at unity power-factor. The coefficient k_2 depends on the power-factor error, and at power frequencies may usually be represented by one or two values for each frequency. The coefficient k_3 at power frequencies is either zero or equal to the conductance of the voltage circuit, according to the method of wattmeter connection, while the coefficient k_4 is similarly either zero or equal to the resistance of the current circuit.

This simplicity is not maintained as the frequency is raised. The coefficients k_3 and k_4 are then affected significantly by the inductances, capacitances and eddy currents associated with the wattmeter and also by the mutual inductance between the fixed and moving coils. The mutual inductance varies with instrument deflection and a corresponding variation in the coefficients k_3 and k_4 occurs. Furthermore, the coefficients may have significantly different frequency-dependent variations according to the method of wattmeter connection on real load, and both sets of values may differ from those on artificial load.

The application of eqn. (5) in such circumstances would be a laborious matter, and it is considered that the upper frequency limit of a dynamometer should be assumed to have been reached when the frequency-dependent part of either k_3 or k_4 affects the value of the wattmeter indication to a significant extent.

The value of the coefficient k_2 is also affected by the mutual inductance between coils, and a test should be made to ascertain whether the effect is significant.

The errors and calibration of a wattmeter will be considered in greater detail after a review of the methods of avoiding errors from stray-capacitance and eddy currents.

(2) UNSCREENED LOADS

In many cases it is not practicable to screen a load and in other cases it is unnecessary, but the possible error arising from the absence of a screen should be assessed as accurately as possible. The first step is to assign a boundary to the load. This boundary need not have any physical existence and should lie entirely in dielectric material or in air. Apertures in the boundary will be unavoidable and will be required for the load terminals, and possibly for the means of transmitting non-electrical power beyond the boundary of the load. The wattmeter should measure all electrical energy consumed within the boundary of the load. This means that there should be no transmission of electrical power across the boundary of the load in either direction. This can only be ensured with certainty for unscreened loads by moving the load boundary to infinity. Measurement of power in unscreened loads with a finite load boundary is therefore an approximate measurement. Such measurements, however, are often made, and it is desirable to estimate so far as possible whether the error arising from power transmitted across the assigned load boundary is likely to be significant.

(2.1) Electromagnetic Coupling

Electromagnetic coupling is likely to be a source of error in the measurement of power in low-impedance loads located near large masses of conducting material. Since the coupling falls off rapidly as the distance between the load and external conductors is increased,* an indication of its magnitude may often be obtained by varying the distance between the load and the nearest large mass of conducting material and observing the resulting change in power indicated by the wattmeter. Electromagnetic coupling is seldom a serious problem when the position of the load can be varied, but if the load position is fixed the only way of assessing the coupling effect may be by calculation from the dimensions of the load and its surroundings. In some cases it may be reasonable to consider the losses in the surrounding conducting material as part of the energy consumed in the load.

(2.2) Admittance Current

Admittance current between a load and its surroundings, either on account of leakage resistance or electrostatic coupling, is often a source of error in the measurement of power in high-impedance loads. A significant difference of current through the two load terminals is an indication of admittance current, but the possibility of admittance current flowing to one part of the load from a source of higher potential and other admittance current flowing from another part of the load to a sink at lower potential has to be borne in mind. Such double admittance currents would not necessarily be revealed by a difference of current at the load terminals. On the other hand, heavy admittance currents might flow between the load and its surroundings without any significant power loss. Experience is usually the best guide as to whether errors are likely to be significant. In many cases the only significant admittance current is to earth or to surroundings at one potential, and if the potential of the load is wholly above or wholly below that of its surroundings, admittance current in one direction only at any instant may be assumed. In such cases, and provided that the power factor of the load is high, an indication of the error in power measurement on account of admittance may be obtained by making two wattmeter measurements with the current coil first in one load line and then in the other line.

Fig. 1 shows diagrammatically a simple case in which the admittance current is assumed concentrated at one point dividing the load into two impedances $R_A + jX_A$ and $R_B + jX_B$. The admittance current is assumed equal to $I(\alpha + j\beta)$ where α and β are pure numbers numerically much less than unity.

* The mutual inductance between two circuits is inversely proportional to the cube of the distance between them at large separations.

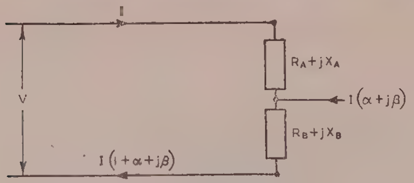


Fig. 1.—Load with leakage current at one point.

The load voltage is

$$V = I(R_A + jX_A) + I(1 + \alpha + j\beta)(R_B + jX_B)$$
$$= I[(R_A + R_B + \alpha R_B - \beta X_B) + j(X_A + X_B + \alpha X_B + \beta R_B)] \quad . \quad (6)$$

The first wattmeter measurement is

$$V.I = I^2(R_A + R_B + \alpha R_B - \beta X_B) \quad . \quad . \quad (7)$$

The second wattmeter measurement is

$$V.I(1 + \alpha + j\beta) \simeq I^2(R_A + R_B + 2\alpha R_B + \alpha R_A + \beta X_A) \quad . \quad (8)$$

The true power consumed in load is approximately

$$I^2(R_A + R_B + 2\alpha R_B) \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Terms in α^2 and β^2 are neglected in eqns. (8) and (9).

If X_A and X_B are zero, the true power lies between the powers indicated by the two measurements. This conclusion also holds if both βX_B and βX_A have the same sign as αR_B. If either or both are of opposite sign to αR_B no conclusion can be drawn.

The following treatment shows that these two wattmeter measurements are also of value with distributed admittance, provided that the load is non-inductive and the active component of the admittance current flows in one direction only at any instant.

Let the current in an element of resistance $d\rho$ at a point separated from one end of the total load resistance, R, by a resistance ρ be $I[1 + f_1(\rho) + jf_2(\rho)]$. $f_1(\rho)$ and $f_2(\rho)$ are pure number functions of ρ, numerically much less than unity for any value of ρ and equal to zero when $\rho = 0$. The function $f_1(\rho)$ has its maximum numerical value when $\rho = R$.

Then

$$V = I\int_0^R [1 + f_1(\rho) + jf_2(\rho)]d\rho$$
$$= RI + I\int_0^R f_1(\rho)d\rho + jI\int_0^R f_2(\rho)d\rho \quad . \quad . \quad (10)$$

The first wattmeter measurement is

$$V.I = RI^2 + I^2\int_0^R f_1(\rho)d\rho \quad . \quad . \quad . \quad . \quad (11)$$

The second wattmeter measurement is

$$V.I[1 + f_1(R) + jf_2(R)] \simeq RI^2 + I^2\int_0^R f_1(\rho)d\rho + RI^2 f_1(R) \quad . \quad . \quad . \quad . \quad (12)$$

The power in the load is approximately

$$RI^2 + I^2\int_0^R 2f_1(\rho)d\rho \quad . \quad . \quad . \quad . \quad (13)$$

Second-order terms are neglected in eqns. (12) and (13).

The power in the load lies between the two powers indicated by the wattmeter measurements since $RI^2 f_1(R)$ is greater numerically than $I^2\int_0^R f_1(\rho)d\rho$.

(3) SCREENED LOADS

The perfectly conducting screen, whether magnetic or non-magnetic, prevents all direct coupling, either electromagnetic or electrostatic, between the load inside the screen and the surroundings, maintains an equipotential surface and is free from loss. If it is also desired to prevent current flowing from the load screen to the surroundings, the surroundings must also be screened at the same potential, although, if the load screen is earth potential, it is necessary to screen only those parts of the surroundings which are not at earth potential. An alternative is to employ a double screen for the load.

Practical screens fall short of perfection, partly on account of their finite conductance and partly on account of the necessary openings in the screens for the admission of leads and for other purposes. Even the perfect screen may be undesirable, on account of the admittance currents between it and the load which are usually appreciably greater than those which would exist in the absence of a screen.

The simplest screened load to consider is that in which the screen and one load terminal are at earth potential, as shown in Fig. 2. If the surroundings are mainly at earth potential there

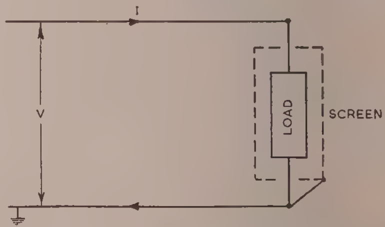


Fig. 2.—Load with single screen at earth potential.

will be little admittance current between them and the screen, but a small amount of current is not objectionable provided that the resistance losses in the screen arising from it are not significant. The proper connection of the wattmeter is with the current coil in the high-voltage line, since the low-voltage line may be carrying a current which has reached the screen from outside. The power indicated will include any loss arising from admittance currents between the load and the screen. This loss may often be reduced to a negligible amount, so far as the dielectric is concerned, by the employment of high-grade insulation, and so far as resistance loss in the screen itself is concerned, by one of the three methods following:

(a) Increase of separation between load and screen.
This method involves a larger screen, which is not always convenient.

(b) Increase of thickness of screen.
This method involves a more expensive screen, but must be adopted if good electromagnetic screening is required, unless a screen of magnetic material is used. There is little advantage in increasing the thickness beyond the "penetration depth" for the frequency and material concerned.

(c) Reduction of screen material.
If electromagnetic screening is unimportant, eddy-current losses may be reduced by using a screen of the wire-cage type with little loss of efficiency as an electrostatic screen.

Some types of wattmeter will not indicate correctly with the current coil in the high-voltage line, and in such cases two modifications of the screening described may be considered.

First, the load screen may be connected to the high-voltage terminal. This means that this screen must be suitably insulated from earth, which is not always convenient. Furthermore, the admittance current between the screen and its surroundings will usually be greatly increased, and this may affect significantly the resistance losses in the screen. The alternative modification is to connect the current coil between the earth point and the load terminal which is connected to the screen. In this case the screen need only be lightly insulated from earth since the voltage drop in the current coil is usually small. Any current between the screen and its surroundings, however, will pass through the current coil and may affect the wattmeter indication.

The measurement of power in systems in which neither line is at earth potential is more complicated. Fig. 3 shows the

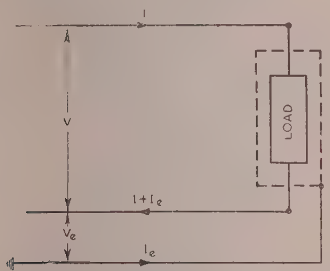


Fig. 3.—Load with neither terminal at earth potential.

$P_L = V.I + V_e.I_e$

method to be adopted when the admittance current between the earthed screen and its surroundings may be considered negligible. The power in the load is determined by two wattmeter measurements in accordance with eqn. (14).

$$\text{Power in load} = V.I + V_e.I_e \quad . \quad . \quad . \quad (14)$$

where V_e is the voltage between earth and the line carrying the current $(I + I_e)$, and I_e is the current between earth and the load screen.

The most severe case to be considered is that in which neither line is at earth potential and the conditions are such that admittance current from the outside to the screen is to be expected. In such cases the load should be fitted with a double screen, the outer one being connected direct to earth. Two wattmeter measurements should be made, one with the current coil in one load line, the voltage circuit between load lines and the inner screen connected to earth, and the other with the current coil connected between the inner screen and earth and the voltage circuit between earth and the other load line. Since the current from the inner screen to earth is usually small, the voltage drop in the current coil will also be small so that the two screens will be nearly at the same potential and there will be a negligible current between them. The conditions are shown in Fig. 4,

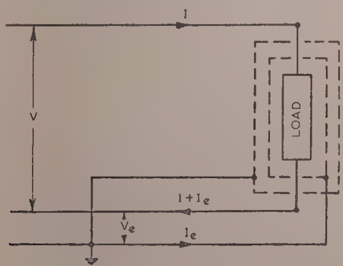


Fig. 4.—Load with double screen.

$P_L = V.I + V_e.I_e$

the power being given by the sum of the two wattmeter readings, as in eqn. (14).

(4) WATTMETER SCREENING

The power-frequency wattmeter is usually fitted with a single screen enclosing the fixed and moving coils. The principal function of this screen is to prevent a torque being exerted on the moving system by external magnetic fields. However, it serves also as an electrostatic screen and is usually connected to a terminal of the moving coil. The series resistor is often unscreened. Assuming that one end of this resistor, of resistance R_s, is at earth potential, the earth capacitance may be represented, as a first approximation, by a capacitance shunting one-third of the resistance; its value might be of the order of 20 $\mu\mu$F. The phase angle of the wattmeter would be altered by such a capacitance by $\frac{1}{9} \times 20\omega R_s \times 10^{-12}$ radn. For $R_s = 10$ kilohms and $f = 1$ kc/s the phase angle change would be 0·0004 radn, which is negligible. The series resistor of a low-voltage wattmeter intended for low audio-frequencies is therefore better unscreened. For $R = 100$ kilohms and $f = 10$ kc/s the phase-angle change would be 0·04 radn. This would introduce a serious error in power measurements at low power-factors. The capacitance between the series resistor and its screen would be greater than the earth capacitance of an unscreened resistor, and screening gives improved wattmeter performance only if compensating capacitance is introduced at the same time.

The coil screen should be connected to the common wattmeter terminal, since, neglecting the effect of mutual inductance between the coils, the potential difference between the screen and the fixed coil is independent of the voltage across the voltage circuit while the potential difference between the moving coil and the screen is independent of the current in the fixed coil. It is convenient for the series-resistor screen to be at the same potential as the coil screen, and the compensation of the capacitance current between screen and resistor is simpler when the screen is connected to the common terminal. It is not therefore proposed to consider any alternative connection of this screen. Fig. 5 shows the method of connecting the wattmeter

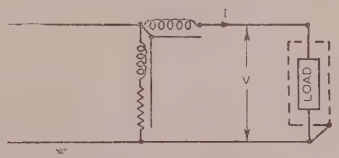


Fig. 5.—Power measurement with singly-screened wattmeter.

to measure a load, one terminal of which is at earth potential. The wattmeter screen is at high potential, but any current between it and the load screen or between it and earth does not affect the measurement except in so far as losses actually occurring in the screens may be altered. The objection to this method of measurement is that the loss in the current coil is included in the power measured, and this is usually more difficult to allow for accurately than the loss in the voltage circuit,

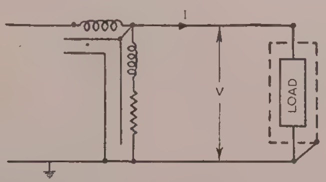


Fig. 6.—Power measurement with doubly-screened wattmeter.

especially at audio frequencies where the effective resistance of the coil may be appreciably higher than with direct current.

This difficulty may be avoided by the employment of a double screen, as shown in Fig. 6. The inner screen is connected to the common terminal and the outer screen to the voltage terminal. The screens form a capacitance in parallel with the load, but if the insulation is of good quality there need be no significant loss associated with it.

(5) THE FREQUENCY ERRORS OF A WATTMETER

The frequency errors of an air-cored wattmeter at power frequencies arise chiefly from the phase angle between the load-voltage vector and the moving-coil current vector. With the Mumetal-cored dynamometer the phase angle between the vector of the current in the fixed coils and the vector of flux produced by it may be the dominating factor. These phase angles give rise to the well-known power-factor error.

At higher frequencies a given voltage is unable to maintain the same numerical value of current in the moving coil, so that the instrument has a frequency error at unity power-factor. In addition, there are frequency errors proportional to the square of the load voltage and load current, as indicated by eqn. (5). These errors arise from eddy-current losses, stray-capacitance currents and the effects of self-inductance, mutual inductance and compensating capacitance. It would be possible to construct a fairly accurate circuit approximation and to derive an expression for the performance of the wattmeter in terms of circuit constants. Such an expression would be extremely complicated and of doubtful practical value. There is, however, some advantage in setting up a very simple circuit approximation in order to see more clearly the nature of the errors involved. Fig. 7 shows such an approximation for an air-cored wattmeter.

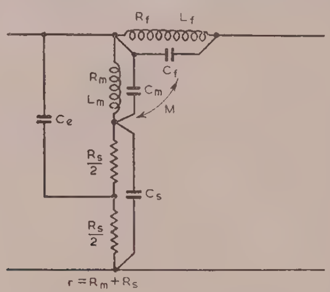


Fig. 7.—Circuit approximation of an air-cored wattmeter.

The effect of eddy currents is very simply allowed for by taking effective values of resistance and inductance of the fixed and moving coils. Eddy-current effects are usually of importance at fairly low frequencies, so that they cannot be neglected. Two capacitances are shown to take account not only of stray capacitances of an unscreened series resistor, but also of capacitance to a screen if one is fitted and of any capacitance added for compensation purposes. The self-capacitances of the fixed and moving coils do not usually have any significant effect at low audio-frequencies. They are shown in the diagram only because they finally introduce errors at the higher frequencies which tend to make the wattmeter unusable. However, troublesome second-order errors would arise at higher frequencies even if these capacitances were negligibly small. The following equations for the wattmeter deflection to a first order of approximation may be derived from the circuit shown in Fig. 7, assuming that the impedances of the fixed and moving coils are small compared to the resistance of the series resistor and that the admittances of the capacitances are small compared to the conductance of the series resistor.

(5.1) Wattmeter on Real Load with the Current Coil Adjacent to the Load

$$P_W = I_m r \,.\, I_f = \left(1 + \frac{R'_m - R_m}{r}\right)P_L + I^2 R_f + [(L_m + M)/r + (\tfrac{1}{4}C_e - C_s)r]\omega VI \sin\phi \quad (15)$$

(5.2) Wattmeter on Real Load with the Voltage Circuit Adjacent to the Load

$$P_W = \left(1 + \frac{R'_m - R_m}{r}\right)P_L + \left[1 + \frac{2}{r}(R'_m - R_m)\right]V^2/r + [(L_m + M)/r + (\tfrac{1}{4}C_e - C_s)r]\omega VI \sin\phi \quad . \quad . \quad (16)$$

(5.3) Artificial Load

$$P_W = \left(1 + \frac{R'_m - R_m}{r}\right)P_L + [L_m/r + (\tfrac{1}{4}C_e - C_s)r]\omega VI \sin\phi \quad . \quad . \quad . \quad . \quad (17)$$

(5.4) $I = 0$ (Artificial Load)

$$P_W = -\,\omega^2 C_f M V^2/r \quad . \quad . \quad . \quad . \quad (18)$$

(5.5) $V = 0$ (Artificial Load)

$$P_W = -\,\omega^2 (C_m + C_e) M I^2 r \,. \quad . \quad . \quad . \quad (19)$$

Eqns. (15), (16) and (17) show that the calibration obtained at unity power-factor on artificial load is applicable on real load provided that the appropriate correction is made for the loss in the fixed coils or the loss in the voltage circuit. The power-factor error on real load differs from that on artificial load by an amount depending on the mutual inductance between the coils. The effect may be assumed negligible at power frequencies but at audio frequencies it is desirable to check the smallness of the error by the following simple test. A convenient current is passed through the fixed coils with the voltage circuit short-circuited through an ammeter of negligible resistance. The ratio of the current in the voltage circuit to the current in the fixed coils is approximately equal to $\omega M/r$ at the deflection realized, which should be negligibly small.

Apart from this mutual-inductance effect the power-factor error may be reduced to zero by adjustment of the capacitance C_s to a suitable value. Adjustment of the power-factor error is also possible for Mumetal-cored dynamometers, but in this case it may be necessary to increase C_e.

The wattmeter indications shown by eqns. (18) and (19) should be negligibly small in the working frequency range of the instrument, and this may clearly be checked by simple tests. It seems reasonable to say that the frequency at which these or other second-order errors become significant constitutes the useful upper limit of frequency for the wattmeter. The variation of mutual inductance with deflection would make it very laborious to apply corrections for these errors.

(6) CALIBRATION OF THE STANDARD WATTMETER

The calibration of the electrostatic voltmeter and the determination of its d.c./a.c. sensitivity ratio have been described in an earlier paper.[1] With the voltmeter calibrated it is possible to calibrate the electrostatic wattmeter with alternating current, using the voltmeter as a standard of voltage and suitable combinations of resistance, capacitance and inductance as the load. The effective values of these components with alternating current must also be known. Resistances must be known to the full accuracy required for power measurements and the design of suitable resistance standards has been

described.[2] Capacitors and inductors are usually used for tests at low power-factor where the wattmeter deflection is small. It is not necessary to know their values to high accuracy, but their phase defect must be accurately known. This practically excludes the use of inductors for real-load tests.

Eqn. (5) may be used as a basis for investigating the errors of an electrostatic wattmeter as well as a dynamometer. One of the great merits of the electrostatic wattmeter is that the coefficient k_2 is zero, at least up to 20 kc/s, so that the instrument has no power-factor error.

The simplest way of verifying this at power frequencies is by the use of mutual inductance, as shown in Fig. 8. The mutual

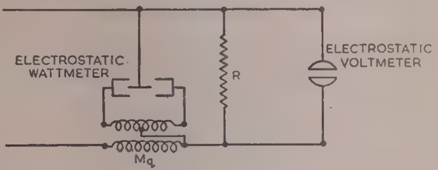


Fig. 8.—Zero-power-factor test of electrostatic wattmeter at power frequencies.

inductance, M_q, is of such a value that when the current through the resistor R gives a voltage drop equal to the full-scale "needle" voltage, the same current through the primary of the mutual inductance induces full-scale "quadrant" voltage in the secondary. If the phase angle of the resistance and the phase defect of the mutual inductance are negligibly small the needle and quadrant voltages will be in quadrature and the wattmeter will give zero deflection provided that it is free from power-factor error. The effect of either leading or lagging power-factor may be obtained by the choice of a suitable polarity for the mutual-inductance connections. It may easily be verified that this test is satisfactory, even for a non-sinusoidal current, provided that there are no even harmonics. At higher frequencies it is difficult to estimate accurately the phase defect of the mutual inductance, and tests with capacitances become easier. A test at a power factor approximating to zero, current leading, may be made as a real-load test with a capacitance load: a similar test at a power factor approximating to zero, current lagging, may be made as a real-load test with a resistance load and a capacitance connected across the quadrants. Both these tests have been carried out at frequencies up to 20 kc/s, and no significant power-factor error has been observed. To complete the power-factor tests the error of the wattmeter was checked at various parts of the scale and at different power factors. It was found that the error at all parts of the scale was independent of power factor. This test was carried out as a real-load test with a load consisting of a known resistance in parallel with a known capacitance.

The current coefficient, k_4, may be determined by applying a voltage between the quadrants with the needle connected to one quadrant. This test was carried out at various frequencies up to 20 kc/s, and it was found that the wattmeter deflection was independent of frequency and proportional to the square of the voltage between the quadrants. The coefficient k_4 is therefore a constant, and its effect is usually allowed for by assuming that half the power dissipated in the quadrant resistance is included in the power measured by the wattmeter in real-load measurements.

Theoretically the voltage coefficient, k_3, should be zero, and it is possible to make it so for one position of the needle by adjustment of the quadrant positions. It is found, however, that exact adjustment is not maintained for long, and it is usually considered preferable to make a correction for a small value of the coefficient than to adjust the instrument continually. By applying a voltage to the needle with the quadrants short-circuited the value of the coefficient near zero deflection may be determined. The needle is then at the "electrical zero," and no other correction need be made if later deflections are measured from this "zero" instead of the mechanical zero. There is the possibility that the coefficient will have a different value at other deflections, but no error need arise on this account if the calibration is determined with the same voltage applied to the needle as the voltage applied when the wattmeter is used as a standard. Finally, by calibrations at different frequencies it has been established that the coefficients k_1 and k_3 are independent of frequency so that it is legitimate to calibrate the wattmeter on one frequency and use it on another frequency. This is convenient for normal testing, as the original wattmeter calibrating equipment described in another paper[3] is not designed for audio frequencies but is more rapid in use than the special equipment described here. The normal procedure is therefore to calibrate the wattmeter at 50 c/s, or a low audio frequency, even when the subsequent test is at a high audio frequency.

As a result of these various tests it has been established that the wattmeter calibration is independent of frequency to one division deflection or 0·05% of full-scale deflection. To calibrate other wattmeters with the electrostatic instrument as a standard, a range of quadrant resistances and a voltage divider must be available to cover the desired ranges of current and voltage. These auxiliaries must have an accuracy comparable to that of the wattmeter, and the voltage divider should be compensated to allow for the shunting effect of the capacitances of the electrostatic voltmeter and wattmeter.

A set of quadrant resistors for currents from 10 milliamp to 1 amp and a voltage divider for voltages up to 1 100 volts have been constructed using the principles of design described in another paper.[2] For currents heavier than 1 amp, existing power-frequency quadrant resistors will be used, a special determination of their phase angle being made.

(7) THE ARTIFICIAL-LOAD METHOD OF TESTING DYNAMOMETERS

The connections for testing dynamometers by artificial load at power frequencies are shown in Fig. 9. Investigation has

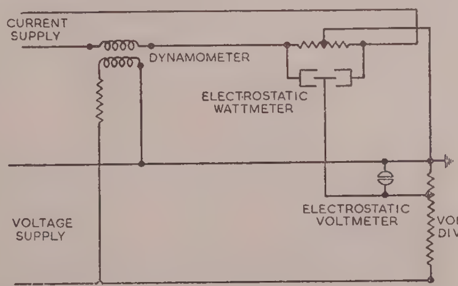


Fig. 9.—Artificial-load test of dynamometer at power frequency.

shown that, provided the test circuits are connected to the power supplies through isolating transformers, the current through the earth connection is not sufficient to affect the accuracy of the tests unless the test current is very small. A similar immunity from error cannot be assumed at audio frequencies, and it is desirable that the earth connection should be made to one supply terminal of each supply, which then forms the common point. The mid-point terminal of the quadrant resistor cannot therefore be the common terminal of the two supplies, so that the necessity for making a correction for half the power dissipated

in the quadrant resistor has to be accepted. Apart from this, the only difference between the audio-frequency test and the power-frequency test is in the precautions taken to avoid error from the capacitance currents taken by the electrostatic instruments and stray-capacitance currents. Fig. 10 shows the connections for audio-frequency tests with the important capacitances indicated by broken lines. The voltage divider is compensated for the capacitance currents taken by the voltmeter and wattmeter as described in another paper.[1] The voltmeter and wattmeter cases and the screens of the voltage divider and the quadrant resistor are all at earth potential. The time-constants of both the voltage divider and the quadrant resistors have been adjusted to zero to an accuracy of $0\cdot002\ \mu H$ per ohm, so that the wattmeter calibration is applicable without correction.

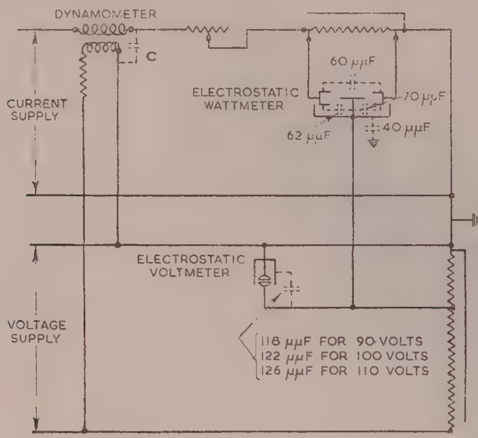


Fig. 10.—Artificial-load test of dynamometer at audio frequency.
Broken lines indicate capacitances between parts.
C = Capacitance, usually unknown, between coils of dynamometer.

The test wattmeter, however, is not tested under exactly the conditions of use. There are two differences to be considered.

(*a*) In use one terminal of the moving coil is connected to a terminal of the fixed coils, while in this test there is a potential difference between the two terminals equal to the voltage drop across the quadrant resistor. It is rare for this difference to affect the wattmeter calibration significantly, but the effect, if any, may be determined by a repeat calibration with the voltage drop increased by the addition of resistance between the quadrant resistor and the wattmeter current-circuit. If the two calibrations are not the same an extrapolation to zero total resistance gives the required figure.

(*b*) The coils of the wattmeter are at or near earth potential, whereas it has been shown in Section 3 that it is often desirable for the current coil to be connected into the high-voltage line. A test for the effect of this may be made by enclosing the wattmeter in a temporary screen connected to the high-voltage line. If the calibration is the same with and without the screen it may be assumed to be reliable, but if there is a difference, the exact size of screen to simulate the conditions of the unscreened wattmeter as normally used is difficult to assess. Under such circumstances there is an unavoidable uncertainty in the calibration of the instrument. If the wattmeter is fitted with a single or double screen it is unnecessary to carry out a test with any added screen. The wattmeter screen, or the inner screen if there are two, must, however, be connected to the moving-coil terminal—and not the fixed-coil terminal—for the artificial-load test, in order that the current from screen to earth shall not affect the measurements.

(8) DESIGN

Little has been said in the paper about the difficulties of internal design of dynamometers for audio frequencies, although the subject was discussed briefly in an earlier paper.[4]

The first difficulty likely to be encountered is in maintaining the current in the moving coil in the correct proportion to the applied voltage and in phase with it. Capacitors may be used to improve the performance in this respect, but the variable effect of mutual inductance prevents perfect compensation of phase at low audio frequencies; at the higher audio frequencies, when second-order errors become appreciable, the ratio of moving-coil current to applied voltage will change with frequency even in the position of zero mutual inductance. There are two possible lines of progress here: the first is to maintain the correct moving-coil current by electronic means[4] and the second is to design the wattmeter to operate with the coils always in the position of zero mutual inductance. This may be achieved by a torsion-head control or by an electrical control as in the Hawkes–Shotter comparator.[5] Such instruments, however, are not so rapid in operation as a deflecting instrument. This difficulty may be overcome by the use of photocells: it is then possible to apply the electrical control automatically, and the value of the control current, read on a deflecting d.c. instrument, provides the indication of power.[6]

The second difficulty in the design of audio-frequency wattmeters arises from the higher impedance of the current circuit in the upper range of frequencies. The natural solution of this difficulty appears to be in the provision of electronic means for maintaining the current in the fixed coils with the application of only a small voltage at the current terminals. The permissible voltage is much less than is available at the voltage terminals so that the problem is to that extent more difficult of solution. Mutual inductance between the coils also increases the difficulties of design, and instruments operating always in the position of zero mutual inductance have a great advantage. A dynamometer wattmeter with electronic control applied both to the voltage coil and to the current coil has been designed and constructed for operation at the single frequency of 10 kc/s.[8]

The third difficulty arises in the upper audio-frequency range from the shunting effect of the coil capacitances, and may be especially acute where the mutual inductance between coils is large. This difficulty may be met by designing the coil systems for best performance and providing auxiliary means to adapt the resulting instrument. With existing designs of dynamometers having pivoted movements the best performance up to frequencies of 20 kc/s appears to be obtained when the coil current is about 30–50 mA with the Mumetal-cored type and 100–150 mA with the air-cored type. The current range of such instruments may be extended by means of current transformers: it may be advantageous to use separate transformers for the upper- and lower-frequency ranges. If, however, electronic control is provided on the current side, the range can be changed by different values of resistance used to provide a suitable voltage for the electronic amplifier. Electronic equipment is likely to be essential on the voltage side. Since a current of 30–50 mA is easier to handle electronically than 100–150 mA, the Mumetal-cored type of dynamometer has an advantage over the air-cored type for the higher audio-frequencies. Lower values of current are likely to be appropriate for sensitive instruments with suspended movements and either the air-cored or the Mumetal-cored type might then be suitable for electronic control. The complete equipment consisting of a dynamometer, electronic amplifiers for the voltage and current side, a voltage divider and a set of resistance standards or current transformers makes a formidable collection: the alternative possibility visualized by Choudhury and Glynne[7] of developing a robust form of the electrostatic wattmeter with electronic control is not without attractions, and it will be interesting to see which type of instrument makes the most progress in the future.

(9) ACKNOWLEDGMENTS

The work described has been carried out as part of the research ogramme of the National Physical Laboratory, and the paper published by permission of the Director of the Laboratory. ne author desires to acknowledge the assistance rendered by essrs. J. J. Hill and A. P. Miller.

(10) REFERENCES

) Arnold, A. H. M.: "Alternating-Current-Instrument Testing Equipment," *Proceedings I.E.E.*, Paper No. 1532 M, July, 1954 (**101**, Part II, p. 121).

) Arnold, A. H. M.: "Alternating-Current Resistance Standards," *Proceedings I.E.E.*, Paper No. 1512 M, June, 1953 (**100**, Part II, p. 319).

) Paterson, C. C., Rayner, E. H., and Kinnes, A.: "The Use of the Electrostatic Method for the Measurement of Power," *Journal I.E.E.*, 1913, **51**, p. 294.

(4) Arnold, A. H. M.: "Performance Limits in Electrical Instruments," *Proceedings I.E.E.*, Paper No. 1128 M, March, 1951 (**98**, Part II, p. 701).

(5) Shotter, G. F., and Hawkes, H. D.: "A Precision A.C./D.C. Comparator for Power and Voltage Measurements," *Journal I.E.E.*, 1946, **93**, Part II, p. 314.

(6) Merz, L.: "Electric Torque Compensators for Remote Measurement and Regulation Technique," *Microtecnic*, 1952, **6**, p. 287.

(7) Choudhury, J. K., and Glynne, A.: "The Use of an Electrostatic Wattmeter for Magnetic Loss Measurements," *Proceedings I.E.E.*, Paper No. 1102 M, January 1951 (**98**, Part II, p. 539).

(8) Boast, W. B.: "An Electronic Wattmeter," *Proceedings of the National Electronics Conference* (*Chicago*), 1952, **8**, p. 716.

DISCUSSION ON THE PAPERS BY PROFESSOR BARLOW AND DR. ARNOLD, BEFORE THE MEASUREMENTS SECTION, 23RD NOVEMBER, 1954

Prof. R. O. Kapp: If I do not refer to Dr. Arnold's paper it is nly because it deals with a subject rather remote from my own eld of study. The use to which Professor Barlow is putting the Iall effect may, on the other hand, have wide applications in eavy current engineering. The advantages of this device from he power engineer's point of view are four. It is nearly freuency independent over a wide range. It has no moving parts. : is free from magnetic saturation up to a practically unlimited eld strength. It occupies very little space. A disadvantage is nat the signal received from the Hall crystal will nearly always ave to be amplified. But I think that we power engineers are nore afraid of amplifying devices than we ought to be in these ays. We have not taken enough notice of the improved techiques that have recently been developed.

One field of application would seem to be in the electrohemical and metallurgical industries where the currents run nto tens of kiloamperes and ordinary shunts prove inadequate. Iere Prof. Barlow's use of the Hall crystal may become a uccessful competitor to the direct-current transformer. In a.c. vork the device would have to be adapted to become a VAr-meter s well as a wattmeter, but this should not prove difficult for it vould suffice to provide one of the elements with a quadrature omponent of the current or voltage in the main circuit. Other ossible uses may be as an ammeter or a voltmeter in cases vhere the frequency independence, the probable reduction in vear and maintenance and the small dimensions are important; or if the device can measure the product of voltage and current t can, of course, also be connected to measure the product of urrent and current or of voltage and voltage. The characeristic would be parabolic and not linear, but there are occasions vhen that too would be an advantage.

Lastly, I should like to mention a further field that I think ught to be explored, even though I dare not hazard a guess vhether the exploration would lead to a fruitful result or not. am thinking of the use of Prof. Barlow's Hall crystals in place f current transformers in very-high-voltage bushings. In ower work it is often necessary to accommodate three sets—ne for metering, another for indicating instruments and other ervices, and the third for operating protective relays. The urrent transformers have to be of the bar-primary type and ake up a great deal of room. The design problem is not at all imple. If Hall crystals could be used instead, a lot of space vould be saved. They would have to be connected in cascade nd located near enough to the conductor to be subjected to a trong field. The current element would have to be taken from oltage transformers. The fact that there was no saturation would make such a device particularly suitable for protection, where Mumetal cannot be used on account of its saturation characteristic. There would be many difficulties, I must admit. Insulation is one, for it may not be easy to provide adequate insulation between the Hall crystals and the conductor and yet have a sufficiently intense magnetic field. It may be also difficult to bring conductors from a voltage transformer to the crystals, and it may be difficult to accommodate an amplifying device immediately adjoining a bushing. I think this would be necessary since any length of lead between the crystals and the amplifier would introduce a source of inaccuracy. But I do not think one ought to regard such difficulties as insuperable. One ought at least to make a good try to overcome them.

Mr. H. D. Hawkes: The question of load screening may be part of circuit design and not instrument design: an ideal instrument should measure power in that circuit without disturbing it in any way.

Admittance currents may be part of the load to be measured and, again, appreciable losses due to the measuring instrument may so upset circuit conditions that any attempt to apply correction would be very laborious, particularly if the circuit had non-linear characteristics.

In general, the higher the sensitivity of the instrument the lower are its losses and the lower are its frequency and power-factor errors.

Dr. Arnold stresses the error caused by mutual inductance at the higher frequencies, and I agree that the ideal dynamometer is a null-point instrument, but the next best is the suspended reflecting type of indicator, where the sensitivity is high and the angular movement of the moving coil and hence the mutual inductance troubles are greatly reduced by using optical amplification.

A number of network analysers are being built in this country for simulating power-network conditions by medium-frequency and low powers. Consider an example where the frequency is 1 kc/s, the nominal voltage level 25 volts and the nominal current 50 mA; measurements of voltage, current, power and reactive power are required anywhere within the network to an accuracy of within $0 \cdot 5\%$ of nominal values.

To do this the sensitivity of the apparatus must be sufficiently high for the voltage side to take less than 50 microamp and the current side to have a voltage drop of less than 25 mV.

This is a specific case for the use of amplifiers for both voltage and current sides of the apparatus.

The instruments we have used for this purpose are the suspended reflecting type of dynamometer with the coils at earth

potential, the primary measuring components being small screened transformers for both current and voltage.

The current side is supplied by resistance-drop feeding a screened transformer and isolating the instruments from the network; the only disturbance of circuit conditions is due to very small screen currents when the resistance potential is above earth in the network. In connection with Prof. Barlow's Hall-effect low-frequency wattmeter, I should like to compare the performance with that of a reflecting dynamometer of similar range. The sensitivity and accuracy comparisons are as shown in Table A.

Table A

Property	Hall wattmeter	Dynamometer
VA on current side	1	0·25
Ω/v of voltage side	60	800
Frequency error, % . .	7 (50–150c/s)	0·2 (0–2 500c/s)
Power-factor error at 0·5 lag, %	—	<0·5 F.S.D.
Torque of indicator, g-cm* . .	0·01 (approximately)	0·03

* The Hall wattmeter indicator is a millivoltmeter.

Dr. E. Billig: Only in the last few years has it become practicable to utilize the Hall effect, since no suitable materials have been available before. The application of the effect is, of course, not restricted to the measurement of power, but can be used to measure the product of any two independent variables as in computers, in the measurement of the torque of a rotating electrical machine (torque = magnetic flux × current), etc.

In the design of Hall plates, the choice of an appropriate material is all-important. Eqn. (2) of the paper gives an expression for the Hall constant in terms of the concentration of electrons (n_e) and holes (n_h). Utilizing the additional condition that their product is constant (for a given temperature and type of semi-conductor):

$$n_e n_h = n_i^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (A)$$

the Hall constant can be expressed as a function of the electron concentration alone:

$$\mathscr{R} = -\frac{3\pi}{8en_i} r(\nu) \quad . \quad . \quad . \quad . \quad . \quad . \quad (B)$$

where $$r \equiv \frac{\nu - (\beta^2\nu)^{-1}}{\nu + (\beta\nu)^{-1}}, \nu \equiv \frac{n_e}{n_i}, \beta \equiv \frac{b_e}{b_h} \quad . \quad . \quad . \quad (C)$$

and n_i = concentration of intrinsic electron-hole pairs.

The factor r has been plotted in Fig. A as a function of ν for various parameters β. This mobility ratio β has a value of 2·1 for germanium and 5 for silicon. It is seen that the Hall constant is highest for near-intrinsic semi-conductors ($n_e \sim n_i$). Plates made of such material would generate a large voltage as long as no load was drawn from the transverse terminals, but relatively little power could be drawn from such a device owing to its rather high internal resistance. In practice, a compromise solution is therefore adopted by using "doped" material, i.e. one in which the number of free electrons is increased by the addition of small quantities of appropriate "donor impurities," such as arsenic, antimony, bismuth, etc. The conductivity of the semi-conductor is thereby increased—at the expense of the Hall constant; i.e. the internal resistance is reduced at the cost of reduced sensitivity, and rather complicated detectors have to be used, such as the photo-electric galvanometer amplifier described in the paper.

A better solution, which has recently become feasible as a result of fundamental research into semi-conductors, is to use alternative materials instead of germanium. The "intermetallic" compounds of Groups III and V, i.e. stoichiometric combinations of an element in Group III of the Periodic Table (aluminium, gallium, indium) with any element in Group V (phosphorus, arsenic, antimony, bismuth) are semi-conductors having properties remarkably like those in Group IV, such as silicon or germanium; in particular, their Hall constants are quite comparable. In addition, however, the mobility of the electrons in these compounds is much higher; in InSb, for instance, by a factor of 20. It thus becomes possible to combine large Hall constant with large conductivity, i.e. to produce Hall plates of low impedance and yet giving a relatively strong signal.

In connection with eqn. (2), I suggest omitting the factor 10^8, and also the rather awkward dimensions given. The true dimension of the Hall constant is that of the volume of semi-conductor occupied by unit free charge. This is independent of the particular system chosen; in practical units, for instance, this would be expressed as cubic centimetres per coulomb.

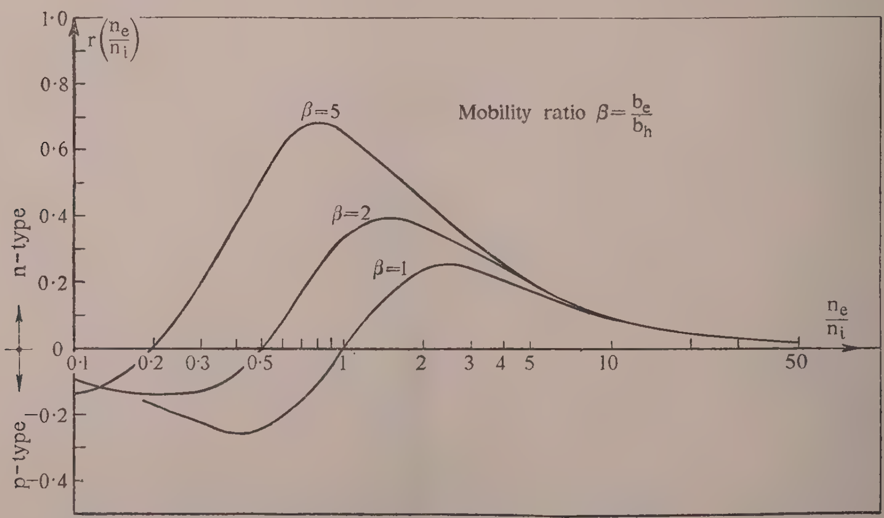


Fig. A.—Hall constant of near-intrinsic semi-conductors.

$$R = \frac{3\pi}{8en_i} r\left(\frac{n_e}{n_i}\right)$$

Dr. A. L. Cullen: I propose to confine my remarks to Prof. rlow's paper on the Hall effect, in which he pointed out the fficulty of measuring the Hall constant at very high frequencies d mentioned the possibility of using for this purpose wave-opagation characteristics in a semi-conductor situated in a eady magnetic field directed along the path of the wave. An itially linearly-polarized wave acquires an elliptic polarization it progresses, and the degree of ellipticity and the orientation the ellipse slowly change as the wave passes into the semi-onductor.

Analysis shows that the attenuation of a TEM wave from the ce of the semi-conductor to the point at which it has acquired n ellipticity such that the ratio of major to minor axis becomes 17 : 1 is a direct measure of the Hall constant if the conduc-vity and the steady magnetic field are known. Putting in gures for a sample of germanium which Prof. Barlow has used n some of his experiments leads to the result that with a wave-ngth of 3cm this degree of ellipticity is reached only after an ttenuation of the order of 50 000dB. Thus, it does not seem hat there is much possibility of measuring the Hall constant in his way.

However, there is a possibility of measuring the rotation of he major axis as the wave progresses, and in that case the ituation is far more favourable. A rotation of half a degree, which should be observable, will take place at a wavelength of cm after an attenuation of $8 \cdot 7$dB. In all cases I am assuming he same magnetic field of $0 \cdot 1$ weber/m^2, or 1 kilogauss.

A final remark concerns the relationship between radiation pressure and the Hall effect. Consider a plain wave incident normally on a conducting metal surface. The usual explanation of the resulting radiation pressure is that the r.f. currents flowing n the surface of the metal are acted upon by the r.f. magnetic ield in which they are situated, giving rise to an inward force which can be calculated from the usual *BlI* rule. This physical mechanism is the same as that which gives rise to the Hall effect. A point which this theory of radiation pressure leaves unexplained is how this force, which acts only on the free electrons, is trans-ferred to the bulk metal. A purely classical picture of this transfer is that the force acting on the electrons tends to push them inside the metal, thereby leaving "uncovered" a layer of immobile positive charges near the surface. This displacement of the electrons inside the metal produces an electric field pointing inwards and tending to pull these positive charges in to follow the electrons.

If one calculates from the known magnitude of the radiation pressure the value of this electric field, it turns out to be exactly the same as the Hall e.m.f. per unit length calculated with the classical result that the Hall coefficient $\mathscr{R}$ is equal to $(ne)^{-1}$ where e is the electronic charge and n is the number of free electrons per unit volume of the metal. Thus, there is a close connection between radiation pressure and Hall e.m.f., and one might say that it is existence of the Hall effect which makes radiation pressure observable.

Mr. A. Hobson: An instrument using a new basic principle is a rare event in the well-established art of electrical science, and Prof. Barlow's device has many interesting possibilities. At the moment it seems that the most likely practical applications will be in measurements on direct currents and at very high frequencies.

At power frequencies it is difficult, at this stage, to see what can be gained by avoiding the use of current transformers. The arrangement shown in Fig. 8 is eminently suitable for direct current, but for a.c. work the core could well belong to a current transformer. It is no easy matter to arrange for the flux density in a fairly long air-gap to have an accurately known value, and for this reason a complete transformer might cost no more than the special core in the diagram.

It has been suggested that a development of Prof. Barlow's instrument might replace the protection-current transformer, and this also I find difficult to imagine. The present protection transformer is big and clumsy because it is called upon to provide considerable power. If the relay circuits could be designed to operate from a Hall e.m.f. of a few millivolts an air-cored toroid would give a bigger signal in a simpler way with complete stability and indefinite linearity.

A very important advantage of the author's device is its inde-pendence of frequency. Measurements of current at very high frequencies are difficult, and the possibility of calibrating an instrument at 50c/s and then being able to use it with confidence at 300Mc/s is an extremely attractive one.

Dr. E. W. Saker: Dr. Billig has mentioned that some of the Groups III–V compounds would be more efficient than ger-manium.

One can show very simply that in devices using the Hall effect the efficiency—i.e. the ratio of Hall power produced in a matched measuring meter to the power dissipated in the specimen—is proportional to the square of the carrier mobility in the semi-conductor. Since the mobility of electrons in indium antimonide is 15 times greater than in germanium, the efficiency could theoretically be over 200 times greater, provided that no practical difficulties arose. Another advantage of indium antimonide is that the contacts are linear.

Mr. M. Kaufmann: I was interested in Prof. Kapp's reflections on the possibilities of using the Hall effect in various devices in the future, and in the reactions which they produced in Mr. Hobson. Although I cannot see any immediate prospect of the Hall-effect device replacing the current transformer, I consider it quite safe to predict that the future will inevitably bring about such changes as will be tantamount to the abolition of the current transformer as we know it now. It is not that the transformer is failing in the role it is required to fill; it is merely that it has definite shortcomings, especially in balanced protective systems.

I agree with Prof. Kapp that protective devices may be a little extravagant in the matter of the power they use, but that has not been a serious difficulty. One that can be called serious is that which arises from the use of iron in balanced protective systems, and the idea of dispensing with iron, although not new, is very attractive.

The difficulty I see in applying the device which Prof. Barlow has described in place of a current transformer is that of insu-lating it. The replacement of a conventional current trans-former by something called a transducer, which this might perhaps be called, has been tried before, but the problem was one of insulating the transducer from the device it had to operate.

Looking at Fig. 12, I think the problem of insulating this device would be simpler than in the case I had in mind, and I should like Prof. Barlow to comment on that. To me it seems possible to put a similar wheel-spoke arrangement of semi-conductors around an ordinary porcelain insulator forming the shedding of a capacitor bushing, and so derive enough force for an instrument. If that can be done, it is well worth following up, since it could lead to a significant development in the art of protection.

In such an arrangement the magnetic field round the central conductor in the bushing would provide the flux, and a tapping from an outer layer of the capacitor could provide the current. The problems then posed would be how to correct the phase of the capacitance current in the semi-conductor; and whether the strength of the magnetic field would be adequate at a distance of, say, one foot from the conductor.

I should like to consider the possibility of being able to compare the phase of the power at two ends of a power circuit

as a first step in the development of a protective device not dependent on conventional relays or conventional current transformers.

Prof. F. Brailsford: I am interested in these papers in connection with the measurement of iron losses in magnetic laminations at high flux-densities and power frequencies. We have recently succeeded in making measurements up to about 24 kilogauss, but this has been by thermal means, and confirmation of these results by an alternative method would be welcome.

From the point of view of a wattmeter the problem is that of low power-factor (e.g. in the range 0·01–0·001) and also of the presence of harmonics in both the voltage and current circuits.

The Barlow wattmeter might possibly be designed with phase-angle errors small enough to enable such measurements to be made. The direct approach to this appears to be to eliminate all iron, with its non-linear characteristics, from the wattmeter circuits and, employing the circuit of Fig. 10 in the second paper with L_o also removed, to produce the necessary magnetic flux-density in the germanium by an air-cored coil. It then becomes necessary to produce the maximum field-strength in a coil absorbing the minimum power but, fortunately, over only a very small volume. Kapitza dealt with this problem in some of his experiments in which very high magnetic field-strengths were employed. A rough estimate indicates that a field of 100 oersteds could be produced in a coil of inside diameter 1 cm for the expenditure of about 2 watts. If field strengths of this order are adequate such a wattmeter might be suitable for special investigations of the kind mentioned above.

Mr. F. R. Axworthy: Germanium probes using the Hall effect for measuring magnetic field-strengths are very sensitive to temperature changes. Will Prof. Barlow say to what extent his wattmeter is so affected and whether he has found it necessary to introduce some means of temperature compensation?

Dr. Arnold mentions Choudhury and Glynne's paper on electrostatic wattmeters. Their modification of Addenbrooke's adjustable quadrant has made the portable electrostatic wattmeter a practical instrument. It is a simple matter to set up the instrument and the electrical and mechanical zeros can be balanced in a few minutes. With the instrument balanced the voltage error is less than 1% for needle voltages varying between 20 and 150 volts.

A disadvantage of the electrostatic wattmeter in its more robust portable form is the relatively high power consumption of the shunts, which makes it necessary to use amplifiers for some purposes. The design of these amplifiers is a somewhat difficult problem if one wants to work at power factors lower than 0·1, as is often the case.

When used without amplifiers the chief source of uncertainty is the time-constant of the quadrant shunt. Recent papers by Dr. Arnold and Prof. Moullin have been of great help in the design of resistors with short time-constants, but having designed them it is still rather difficult accurately to measure the time-constant. Perhaps Dr. Arnold could recommend a suitable method.

When considering wattmeters for use at audio frequencies, the electrostatic wattmeter is probably easier to design and make than the dynamometer and will almost certainly have a greater frequency range.

Mr. W. Bamford: I was somewhat carried away by Prof. Kapp's views of the future, but, thinking about it afterwards, I realize that, despite the attractiveness of the Hall effect, it does not provide a great deal of power in terms of operation of instruments and meters. Consider, for example, the measurement of heavy direct current. Although the voltage across the Hall plate is some two volts, the output is certainly not enough to operate an ampere-hour meter on direct current. Amplification has therefore to be used, and by the time we have finished we are back to something as bulky as the so-called d.c. current transformer with its separate a.c. supply. This device, incidentally, is very robust and has operated very successfully, over many years, in many installations.

My questions concern the stability of the Hall effect on continuous load and on overload. Is the output quite stable? The effect of ambient temperature has been mentioned earlier. Is the output affected by a magnetic field which is not regular over the surface? Is it also affected by slight changes in its physical position within the magnetic field? These are practical points to be considered when designing an instrument for use on power supply, in normal working conditions.

Mr. C. M. Burrell (*communicated*): In radar work, requirements for accurate power measurements are mainly at the following power levels:

10^5 to 10^6 watts
10^{-3} to 10^{-2} watt
10^{-15} to 10^{-13} watt

It is of some interest, therefore, to estimate the performance of this new power-measuring device at these power levels.

Evidently the output at the lowest power level would be insignificant, as would the output of anything but a good receiver. At the level 10^{-3} to 10^{-2} watt it is perhaps permissible to obtain figures from the busbar experiment.

If the voltage across the crystal is 5 volts and the resistance of the crystal is 650 ohms, we find that the power dissipated in the crystal is 4×10^{-2} watt, the output e.m.f. is 6μV and the output per dissipated milliwatt is $1{\cdot}5 \times 10^{-7}$ volt.

It is conceivable that a practical device could be made in which E/H was 100 times less than in the example given.

In this case we would have $1{\cdot}5 \times 10^{-5}$ volt per dissipated milliwatt. The corresponding figure for a thermistor would be 10^{-1} milliwatt.

For low-power measurements, therefore, the Hall-coefficient method may be difficult to apply unless a considerable improvement in materials is obtained.

At the highest levels there is the advantage that the power is pulsed, and therefore thermal effects will be reduced by a factor of about 1 000. The crystal might then stand a pulsed dissipation of about $10^{-1} \times 10^3$ watts, and would produce an output of the order of $1{\cdot}5 \times 10^{-2}$ volt.

Here the position appears much more favourable, and the usefulness of a device producing a rectified pulse proportional to power flow would be considerable.

THE AUTHORS' REPLIES TO THE ABOVE DISCUSSION

Prof. H. E. M. Barlow (*in reply*): Some idea of the potentialities of the semi-conductor type of wattmeter for measurement of power involving heavy currents, without the use of a current transformer, can be obtained by a simple calculation. Consider a conductor of circular cross-section carrying 1 000 amp: the magnetic flux in air at a radius of 1 ft is about 6·5 gauss, and this would produce in a germanium crystal, 1 mm thick, carrying a current of 25 mA, a Hall e.m.f. of 0·13 mV. It would not therefore be unreasonable to mount such a crystal at the earth side of a high-voltage insulator bushing, as discussed by Prof. Kapp, Mr. Kaufmann and Mr. Hobson, the current through the crystal being obtained by a voltage transformer or even by a leak circuit taken from the high-voltage line. If a neutral conductor carrying the load current were available, or could con-

niently be made available, the crystal might well be placed mediately adjacent to it so that the flux and the Hall e.m.f. uld thereby be increased perhaps ten times. There is also possibility of a cascade arrangement of crystals around the nductor, and in this way, for such heavy currents, amplification the Hall e.m.f. could probably be avoided.

In its application to electrical protection Mr. Hobson seems overlook the fact that this semi-conductor device does far ore than a current transformer. It responds to reversals of wer in the load circuit and that is particularly valuable in some rms of protection.

Drs. Billig and Saker called attention to the possibility of using her semi-conductors, such as indium antimonide, for this rpose, with a view to taking advantage of the larger Hall efficients available in some cases. Whilst I agree that there e probably better materials than germanium, it must be membered that for a given current density and a given electron obility, the drift velocity varies inversely as the conductivity. earing in mind that the Hall effect can be computed as the m.f. induced in a conductor of length l equal to the distance etween the Hall contacts, moving with the drift velocity u at ght angles with the externally applied magnetic field B, it is ear that the advantage gained by a large electron mobility may asily be annulled by a large conductivity.* At the same time is important to ensure that the power dissipated in the semi-onductor element per unit volume (or preferably, perhaps, per nit surface area) is not such as to cause a substantial rise of mperature, and this requirement is more easily satisfied when he conductivity is high. Semi-conducting materials to which on-rectifying contacts can be made without difficulty are also nost helpful in this application.

If an air-core magnetizing coil is used in this wattmeter, robably the most serious factor limiting its accuracy at low ower-factors is stray capacitance—and particularly the self-apacitance—of the coil. At 50c/s, even if a large range of armonics is present, a high degree of accuracy is possible using, f necessary, reactance compensation, and I feel confident in ssuring Prof. Brailsford that a satisfactory instrument could be nade for his purpose, especially if the power to be measured mounts to a watt or two.

As Mr. Axworthy remarked, the magnitude of the Hall effect in germanium varies with temperature. The smaller the impurity content the more noticeable is this variation, particularly in the neighbourhood of room temperature, but with n-type germanium of resistivity about 37ohm-cm, as in the instruments described, no serious inaccuracy was found to arise in that way. Typical curves given by Shockley† for silicon show that over a range of about 10°C around the ambient temperature, for a material of suitable resistivity, the Hall effect is substantially constant. At 70–80°C the Hall effect in germanium practically disappears altogether, but the crystal is found to recover its properties on cooling. Apart from temperature considerations, semi-conductors are often sensitive to light, but a coating of black paint will generally eliminate any difficulties in this respect. Mr. Bamford need have no uneasiness about the stability of the Hall effect in germanium, subject only to the permissible limits of power dissipated in the crystal. The physical position of the crystal in the magnetic field and non-uniform distribution of that field over its surface are factors that are quite immaterial provided that there are no time variations of these quantities, the crystal being rigidly fixed in position.

For high-frequency measurements Mr. Burrell has applied the criterion of output per milliwatt dissipated and whilst this is undoubtedly quite a good basis for comparison, there are other considerations which may, in a particular application, be more important. It is too early yet to give any reliable information about the possibilities of the device for high-frequency work. Microwaves, for example, produce in germanium a displacement current which is comparable with the conduction current and it is clear that with this added complication and others associated with these conditions there is a wide field for exploration.

Dr. A. H. M. Arnold (*in reply*): The use of screened input transformers on both the voltage and current sides of a wattmeter with feedback amplifiers, as advocated by Mr. Hawkes, in order to obtain freedom from limitations with regard to earthing conditions, is attractive. There is, however, some difficulty in obtaining the required performance, combined with low losses, from the transformers, especially if they are required to operate over a considerable frequency range. The design of the current shunt is relatively easy for the case quoted by Mr. Hawkes, namely one to have a voltage drop of 25mV with a current of 50mA and to have a small phase angle at 1kc/s. The design would be much more difficult for a current of say one ampere at a frequency of 20kc/s, and very difficult indeed for a current of 100amp.

The development of a robust portable electrostatic wattmeter, referred to by Mr. Axworthy, is of great interest, since the electrostatic wattmeter is inherently more suitable for operation at high frequencies than the dynamometer, on account of its freedom from power-factor error. Probably the simplest method of calibrating the time-constant of the current shunts is to observe the wattmeter deflection on zero leading power-factor using a capacitor of good quality as the load. If a resistance shunt of known time-constant is available, any error introduced by the amplifier or by losses in the capacitor may be eliminated by repeating the test using the shunt of known value.

I am rather doubtful whether Prof. Brailsford's problem can be solved by using a wattmeter of any type. The best types available at present might be read to an accuracy of about one part in 10 000 of full-scale deflection, and this would give an accuracy of only 10% at a power factor of 0·001. The problem of making the correct allowance for the effect of harmonics in the voltage waveform would be the same for the wattmeter measurement as for the thermal measurement. However, if a capacitor is connected across the terminals of the magnetizing coil the power factor of the power measured by the wattmeter can be increased to any desired extent and the harmonics in the applied voltage wave might also be reduced. The accuracy of the measurement would then probably be limited mainly by the accuracy with which it was possible to determine the losses in the capacitor.

* Since we are concerned with a semi-conductor, a factor $3\pi/8$ appears in the expression for the Hall e.m.f., namely $v = -3\pi/8Blu \times 10^{-8}$ volt.

† SHOCKLEY, W.: "Electrons and Holes in Semi-Conductors" (D. Van Nostrand and Co., 1950, New York and London).

621.376.56

Paper No. 1776 R
Mar. 1955

SIGNAL/NOISE RATIO IN PULSE CODE MODULATION

By N. L. YATES-FISH, M.A., D.Phil., and E. FITCH, B.Sc.

(*The paper was first received* 10*th December*, 1953, *and in revised form*, 9*th October*, 1954.)

SUMMARY

Formulae are derived for the output signal/noise ratio in simple pulse code modulation systems for all values of input signal/noise ratios. The output ratio improves very rapidly with increasing input ratio provided the latter exceeds a certain critical value. It is shown that the system is useful for links connected in tandem, for the degradation of performance below a single link may be avoided by a relatively small increase of power in each link.

LIST OF SYMBOLS

A = Peak amplitude of the modulating wave which may be transmitted.
V = Pulse amplitude in a practical system.
F = Moment generating function (one link).
G = Moment generating function (M links).
M = Number of links in tandem.
P_N = Noise power in reconstituted pulses.
$P_N(f)$ = Noise power as a function of frequency.
P_{Na} = Noise power in the audio band.
P_S = Signal power in reconstituted pulses.
P_{Sa} = Signal power in audio band.
a = Quantized amplitude of a sample.
a_r = Contribution of rth pulse in a code group.
f = Frequency.
f_a = Highest frequency of the audio band.
f_r = Sampling frequency.
l = Pulse length of reconstituted sample.
m = Number of pulses in a code group.
n = Instantaneous noise amplitude.
n_0 = R.M.S. noise amplitude.
p = Probability of error in a pulse.
p_M = Probability of an error after M links.
p_r = Probability of an error in the rth pulse.
s = Difference between levels in proportion to A.
$\left.\begin{matrix} x_r \\ y_r \end{matrix}\right\}$ = Weights of positive and negative values of the rth pulse.

(1) INTRODUCTION

Prior to the introduction of pulse code modulation, all speech communication systems operated on the principle of causing some parameter of the transmitted signal, such as the amplitude or frequency of a carrier wave, to vary continuously with the instantaneous amplitude of the modulating waveform. Pulse modulation systems, it is true, must of necessity be restricted to the transmission of a sequence of samples of the modulating wave taken at discrete intervals of time; but in systems other than pulse code modulation the modulated parameter is a continuous function of the amplitude of the sample which it represents. In other words, the transmission is quantized in time but not in amplitude. In pulse code modulation, however, the amplitude of each sample is transmitted by a code group of pulses which represents it in digital form, usually on a binary number scale, so that on-off or two-position pulses may be used. Since the number of different code groups which may be formed is limited by the number of digit pulses allotted to the code, the amplitude of a sample can only be approximated by the nearest of a finite number of discrete values; the transmission is quantized in amplitude as well as in time.

This system allows some of the techniques of telegraphic communication to be applied to the transmission of continuously varying waveforms, such as those of speech, with consequent advantages. In particular, it is claimed[1] that the use of signal regeneration enables the quality of the received signal to be made independent of noise or distortion in the transmission medium, provided these do not exceed a certain limit, and that the signal can be relayed through an indefinitely large number of intermediate stations using regenerative repeaters without suffering any cumulative degradation or requiring any compensating increase of power in each individual link.

These claims are based on the assumption that there is some threshold value of the received signal strength such that, when it is exceeded, the receiver can distinguish unequivocally between the two possible states of a transmitted digit pulse. This assumption is not strictly correct, since the statistical properties of the random noise that inevitably accompanies the pulse train are such that there is always a finite and calculable probability that a noise peak will reach a sufficient amplitude to cause the receiver to respond incorrectly, thus changing the value of a received sample and producing a noise pulse in the output. It is the intention of the paper to consider this feature of pulse code modulation quantitatively.

(2) PRINCIPLES OF OPERATION OF PULSE CODE MODULATION

The principles of operation of pulse code modulation are illustrated in Fig. 1 for the case of a 3-digit binary code. The total amplitude range occupied by the modulation waveform [Fig. 1(*a*)] may be thought of as divided into the eight equal regions shown, corresponding to the 2^3 possible combinations of three binary digits. The amplitude of the waveform is sampled at a succession of regularly recurring instants t_1, t_2, . . . , and for each sample the amplitude is taken to correspond to the mid-point of the region within which it falls. In the interval between one sample and the next a code group of three pulses is transmitted, representing the amplitude of the sample by a 3-digit binary number [Fig. 1(*b*)]. At the receiver the three pulses are given their appropriate weights in the ratio 4 : 2 : 1 and combined to reconstitute the original sample as nearly as the quantizing approximation will allow. The result is an amplitude-modulated pulse train [Fig. 1(*c*)] which, on being passed through a suitable low-pass filter, yields a signal which approximates to the original modulation. The horizontal dotted lines in Fig. 1(*c*) represent the eight *quantizing levels* of this system, and each corresponds to the centre line of one of the amplitude regions of Fig. 1(*a*).

Since the processes of time quantization (sampling) and amplitude quantization each involve an approximation, it cannot be expected that such a system will be distortionless. The trans-

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Dr. Yates-Fish and Mr. Fitch are at the Signals Research and Development Establishment.

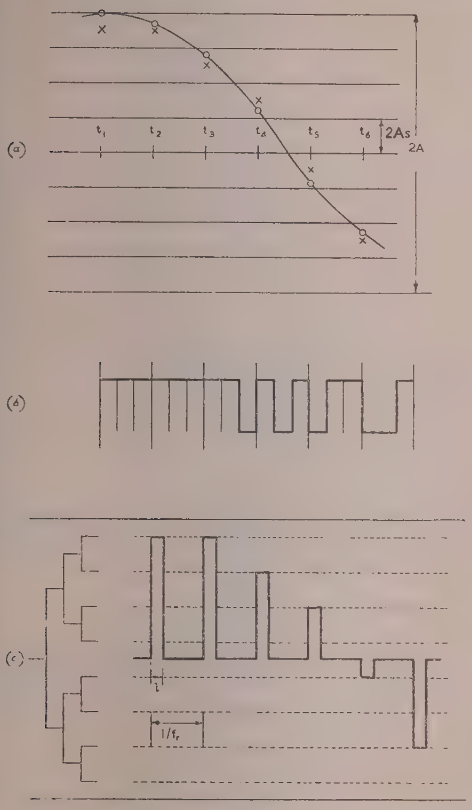


Fig. 1.—Waveforms in pulse code modulation.
(a) Modulating waveform.
(b) Transmitted pulse code.
(c) Reconstituted samples at receiver.
O Sampling points.
× Quantizing approximations.

nission of intelligence by recurrent sampling is common to all pulse modulation systems, and it can be shown that this feature need introduce no distortion provided that the samples are taken t regular intervals and at a rate greater than twice the highest requency component in the modulating waveform. The ampliude quantization process, however, necessarily introduces some listortion in the transmission of a continuously varying wave. This distortion is usually referred to as noise (*quantizing noise*) since, although it is a function of the modulation, it is not harmonically related to it. Quantizing noise has been discussed elsewhere, for example by Bennett.[2] It is difficult to assess, since it depends in a complex fashion upon the precise amplitude and shape of the modulating wave. In principle, it can be reduced below any desired limit by a sufficient increase in the number of digits in the code. The paper is not concerned with his type of noise, apart from noting that in any given system t will set an upper limit to the signal/noise ratio attainable.

It should be mentioned, however, that the effect of quantizing listortion of speech is considerably reduced if the quantized evels are unequally spaced, the size of the quantum step being ess towards the centre of the amplitude range and greater towards he edges. This allows the weaker components of speech to have a relatively greater share of the total number of levels. In practice, the same effect can be obtained by passing the speech wave through an instantaneous waveform compressor before applying it to the coder, and passing the output from the decoder t the receiving end through a waveform expander having a characteristic which is the inverse of that of the compressor. Here, however, it will be assumed that the quantizing levels are equally spaced, since this greatly simplifies the problem without invalidating any of the general conclusions.

(3) THE BASES OF THE METHOD OF CALCULATION

The method to be used assumes that there is some probability that the receiver will respond to a pulse transmitted in one of its two possible states as if it had been in the other state. For example, in an on-off pulse system a pulse which should be present is deemed to be absent and vice versa. It is an essential feature of the system that the receiver be so constructed by the use of trigger circuits or extreme limiting that the possibility of any intermediate type of response is eliminated. The errors are supposed to occur in a random fashion. This is a reasonable assumption for the effect of random noise on the systems so far envisaged, for the correlation function of the noise voltage must be substantially zero at a time corresponding to the interval between adjacent pulses.

The output signal/noise ratio will be calculated when the modulation waveform is a sine wave of as high an amplitude as the system will transmit. There is no difficulty in extending the treatment to other waveforms.

It will be assumed that the code used to represent the quantized samples is binary. Fig. 2(*a*) gives a demonstration of 3-pulse

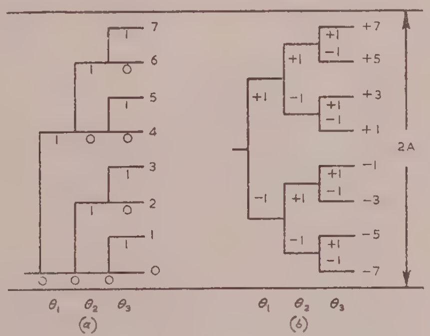


Fig. 2.—Alternative binary representations.
(*a*) Using digits 0 and 1. (*b*) Using digits +1 and −1.

binary code. Let the first pulse of the group have the value θ_1, being either 0 or 1; let the second be θ_2 and the third θ_3; then, if the "tree" is followed, the value of the corresponding quantized sample will be seen to be $4\theta_1 + 2\theta_2 + \theta_3$.

This Figure could be taken as the basis of what follows, but it suffers from the disadvantage that the value of the mean level changes with the number of levels assumed. It is more convenient to base the notation on Fig. 2(*b*). It is assumed that the peak modulation is within $\pm A$. Following the "tree," a pulse train $(\theta_1, \theta_2, \theta_3)$, where each of the θ's can be ± 1, will represent a quantized sample

$$A(\tfrac{1}{2}\theta_1 + \tfrac{1}{4}\theta_2 + \tfrac{1}{8}\theta_3)$$

In general, the quantized amplitude corresponding to an m-pulse code can be written as

$$a = A \sum_{r=1}^{m} \frac{\theta_r}{2^r} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Not only is this notation more convenient from the theoretical viewpoint, it is more closely associated with some practical systems.[3]

It will be convenient to use the symbol s to represent the size of the step of the quantizing process, referred to the signal range,

so that $s = 1/2^m$. The greatest positive value of the quantized amplitude, obtained when $\theta_1 = \theta_2 = \ldots = \theta_m = 1$, is

$$A \sum_{r=1}^{m} \frac{1}{2^r} = A(1 - s) \quad . \quad . \quad . \quad . \quad (2)$$

The maximum error in amplitude occasioned by the quantizing process is $\pm As$. The effect of pulse errors on the train of reconstituted samples applied to the receiver output filter is illustrated in Fig. 3 for a 3-digit code. Fig. 3(a) shows part

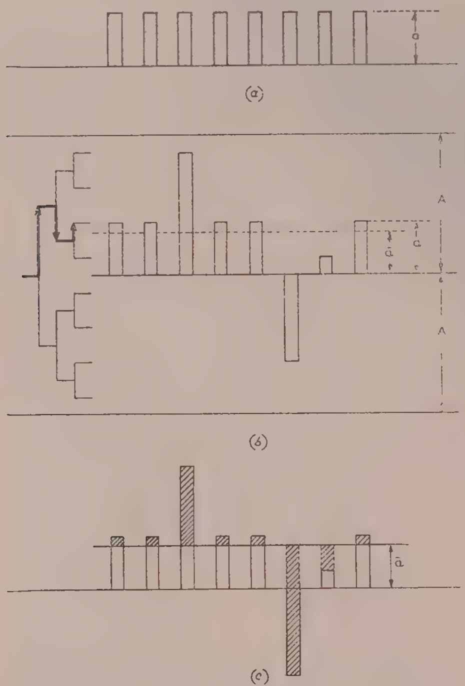


Fig. 3.—Disturbance of a received pulse train by errors.

of a train of equal samples, each of amplitude a, corresponding to the transmitted code group $(\theta_1, \theta_2, \theta_3) = (+1, -1, +1)$. This pulse train, of course, represents modulation by a direct current. Fig. 3(b) shows the corresponding received pulse train in which the code groups producing the samples 3, 6, 7 have suffered errors in the second-, first-, and third-digit pulses respectively. This may be seen by tracing the tree diagram on the left, in which the thick arrows represent the contributions of correctly-received pulses, taking into account the fact that an error in a digit pulse reverses the sense of the corresponding arrow. It will be found that the mean amplitude of the pulses in the disturbed pulse train has a value $\bar{a}$ which is different from a. This new mean value represents the signal component, which in this case has been reduced in amplitude by the errors; the pulse train that remains after this has been subtracted, as shown by the shaded areas in Fig. 3(c), represents noise.

In the general case of modulation by a varying wave, a sample which was transmitted as an amplitude a can be regarded as being received with an amplitude $\bar{a} + n$, where $\bar{a}$ is the mean value of all samples which started as a, and the remainder, n, is the noise component. If the pulses reconstructed in the receiver are rectangular, of length l, and the sampling frequency is f_r, the signal power will be

$$P_S = (\bar{a})^2 l f_r \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

and the noise power will be

$$P_N = \overline{n^2} l f_r \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

It follows that

$$\frac{P_S}{P_N} = \frac{(\bar{a})^2}{\overline{n^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

It is not difficult to show that the power spectrum of noise in a pulse amplitude modulation system has the same distribution as that of a single unmodulated pulse, namely

$$P_N(f) = K\left(\frac{\sin \pi l f}{\pi l f}\right)^2 \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where l is the pulse length and K is constant. Integrating from $f = 0$ to $f = \infty$ and equating the result to the total power P_N, the value of K is found to be $2P_N l$ so that

$$P_N(f) = 2P_N l\left(\frac{\sin \pi l f}{\pi l f}\right)^2 \quad . \quad . \quad . \quad . \quad (7)$$

The noise power falling into the audio band $0 < f < f_a$ is found by integrating this with respect to frequency. Assuming that the pulses are so short that $lf_a \ll 1$, the expression in brackets may be taken as unity and the noise power in the audio band is

$$P_{Na} = 2P_N l f_a \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

It will be convenient to measure the noise relative to a reference signal consisting of a sine wave of the greatest amplitude accepted by the system. If the frequency of this sine wave is f, the power in the audio component is[4]

$$P_{Sa} = P_S l f_r \left(\frac{\sin \pi l f}{\pi l f}\right)^2 \quad . \quad . \quad . \quad . \quad . \quad (9)$$

and if short pulses are assumed

$$P_{Sa} = P_S l f_r \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

The signal/noise ratio in the audio band is consequently

$$\frac{P_{Sa}}{P_{Na}} = \frac{P_S f_r}{2P_N f_a} \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

and if the audio band is half the repetition frequency

$$\frac{P_{Sa}}{P_{Na}} = \frac{P_S}{P_N} = \frac{(\bar{a})^2}{\overline{n^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

The assumption of short pulses is not justified in general. For a given pulse amplitude the amplitude of the audio-frequency output signal is proportional to the pulse length, and so it is convenient to make l as large as possible, namely $1/f_r$. This would give a loss of 4 dB at the higher modulation frequencies, and would require some equalizing network. Such a network would boost the noise in the same way and the final signal/noise ratio would still be as given in eqns. (11) and (12).

In view of eqn. (11) it would seem that to transmit a given band of audio frequencies it would be advantageous to use a higher pulse repetition frequency. This would, however, widen the frequency range of the signal and so lead to an increase of the noise power, which would more than offset the advantage.

(4) APPROXIMATION FOR HIGH SIGNAL/NOISE RATIOS

If the signal/noise ratio is high, the probability that a pulse will be altered is small, and it will be assumed in this Section that the probability is so small that the chance of more than one error in a code group may be neglected.

Suppose that p_r is the probability of an error in the rth pulse a code group. Then for a proportion p_1 of the time the nplitude may change by $\pm A$, for a proportion p_2 by $\pm \frac{1}{2}A$ and on; the mean-square shift is, neglecting the small change in ean value,

$$\overline{n^2} = p_1 A^2 + p_2 \frac{A^2}{4} + \ldots + p_m \frac{A^2}{2^{2m-2}} \quad . \quad . \quad (13)$$

in particular $p_1 = p_2 = \ldots = p$ say,

$$\overline{n^2} = pA^2\left(1 + \tfrac{1}{4} + \ldots + \frac{1}{2^{2m-2}}\right)$$

$$= pA^2 \tfrac{4}{3}(1 - s^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

he important thing to notice is that this noise power is inde-endent of the amplitude or waveform of the signal. This is a onsequence of the assumptions made, in particular of the equal pacing of quantizing levels.

For a given value of p the noise power increases with the number f pulses in a code group, approaching $4pA^3/3$ asymptotically. he convergence is very rapid, as may be seen from Fig. 4,

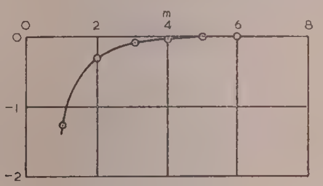


Fig. 4.—Dependence of output noise power on the number m of digit pulses per code group, for constant probability of error.

vhere $1 - s^2$ is plotted on a decibel scale as a function of the umber of digits. The difference from the asymptotic value is ess than 0·1 dB for any number of digits greater than 2 and is herefore negligible for almost all practical purposes. In fact, o matter how many digit pulses are used in the code, the pulse f greatest weight contributes at least three-quarters of the noise ower.

For small values of p the effect of pulse errors on the signal ower may be neglected. The power of a sinusoidal wave of mplitude A is $\frac{1}{2}A^2$, and so the signal/noise ratio is nearly

$$\frac{P_{Sa}}{P_{Na}} = \frac{3}{8p} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

On a decibel scale this becomes

$$10 \log \frac{P_{Sa}}{P_{Na}} = -10 \log p - 4{\cdot}3 \text{ decibels} \quad . \quad . \quad (16)$$

It is well to realize the significance of the numerical values of ignal/noise ratios which may be encountered. It follows from qn. (15) that when the signal/noise power ratio is 1 000 (30dB) he appropriate value of p is 3/8 000. If the sampling frequency s 8kc/s and the number of pulses in a code group is m, there vill be 8 000m pulses transmitted per second, of which $3m$ will e in error. So, on the average, a major error will occur three imes per second. At 40dB there will be a major error about nce in three seconds. Such noise is more closely allied to mpulse noise than to thermal noise, and its annoyance value is ot necessarily the same as that of thermal noise of the same ower. In particular, when the output signal/noise ratio is bove 40dB, the transmission can probably be regarded as being early perfect.

(5) ACCURATE EXPRESSION FOR SIGNAL/NOISE RATIO

(5.1) Signal Suppression due to Noise

In the above Section it was assumed that the noise did not change the mean value of a sample. If the probability of error is not small, this effect cannot be neglected, and it can be evaluated as follows or by a method using probability generating functions[5] as in Appendix 10.1.

Let a_r be the contribution of the rth pulse in a code group to the magnitude a of a sample which is undisturbed by noise, so that

$$a = \sum_{r=1}^{m} a_r \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

If p_r is the probability that the rth pulse will be reversed by noise, its contribution to the output will be a_r for a fraction $(1 - p_r)$ of the time and $-a_r$ for a fraction p_r of the time. The mean value of its contribution is therefore

$$\bar{a}_r = (1 - p_r)a_r + p_r(-a_r) = (1 - 2p_r)a_r \quad . \quad . \quad (18)$$

The mean value $\bar{a}$ of all samples which would have had a value a is consequently

$$\bar{a} = \sum_{r=1}^{m} \bar{a}_r = \sum_{r=1}^{m} (1 - 2p_r)a_r \quad . \quad . \quad . \quad . \quad (19)$$

If the probability of an error has the same value p for all pulses in a code group,

$$\bar{a} = (1 - 2p) \sum_{r=1}^{m} a_r = (1 - 2p)a \quad . \quad . \quad . \quad (20)$$

This means that, if each pulse of a code group has a probability p of being in error, the mean value of the code group is reduced in the ratio 1 to $1 - 2p$. This statement is true for any number of digits. It implies that noise causes a reduction in signal amplitude without altering its form.

It will be noticed that p has not been restricted. For example, if $p = \frac{1}{2}$ the mean signal value is zero; this is to be expected, for then each pulse of a group is equally likely to be positive or negative, no matter what was its original value. If $p = 1$, the mean signal is reversed in sign; this simply means that each pulse is bound to be inverted.

(5.2) Accurate Formula for Noise

To evaluate the mean-square error, it is first noted that the mean-square signal measured from zero in the absence of noise is

$$\overline{a^2} = \left(\sum_{r=1}^{m} a_r\right)^2 = \sum_{r=1}^{m} a_r^2 + \sum_{r=1}^{m} \sum_{\substack{j=1 \\ r \neq j}}^{m} a_r a_j \quad . \quad . \quad (21)$$

The effect of an error is to reverse the affected pulse. This leaves the first term unchanged; terms of the second sum will be reversed if one or other of the factors is changed, but not both. The probability of this is $p_r(1 - p_j) + p_j(1 - p_r)$ when its effect is to subtract a quantity $2a_r a_j$ from the right-hand side of eqn. (21). Consequently, in the presence of noise,

$$\overline{a^2} = \sum_{r=1}^{m} a_r^2 + \sum_{r=1}^{m} \sum_{\substack{j=1 \\ r \neq j}}^{m} a_r a_j [1 - 2p_r(1 - p_j) - 2p_j(1 - p_r)]$$

$$= \sum_{r=1}^{m} a_r^2 + \sum_{r=1}^{m} \sum_{\substack{j=1 \\ r \neq j}}^{m} a_r a_j [1 - 2p_r - 2p_j + 4p_r p_j] \quad . \quad . \quad (22)$$

The square of the mean value in the presence of noise is, from eqn. (19),

$$\bar{a}^2 = \left(\sum_{r=1}^{m}(1-2p_r)a_r\right)^2$$

$$= \sum_{r=1}^{m}(1-2p_r)^2a_r^2 + \sum_{\substack{r=1\\ r\neq j}}^{m}\sum_{j=1}^{m} a_ra_j[1-2p_r-2p_j+4p_rp_j] \quad . \quad . \quad . \quad . \quad (23)$$

The noise component, which is the mean-square deviation about the mean value, is

$$\overline{n^2} = \overline{a^2} - \bar{a}^2$$

$$= \sum_{r=1}^{m} a_r^2[1-(1-2p_r)^2]$$

$$= \sum_{r=1}^{m} a_r^2 4p_r(1-p_r) \quad . \quad . \quad . \quad . \quad . \quad (24)$$

If all the probabilities take the same value p,

$$\overline{n^2} = 4p(1-p)\sum_{r=1}^{m} a_r^2 = \tfrac{4}{3}p(1-p)A^2(1-s^2) \quad . \quad (25)$$

This expression, which applies for all values of p, is more accurate than the previous expression, (14). It will be noticed that the noise power is zero both when $p = 0$ and when $p = 1$, for the latter case implies simple reversal of signal with no randomness. Maximum noise occurs when $p = \frac{1}{2}$, corresponding to a completely random response, and in this case $\overline{n^2} = \frac{1}{3}A^2(1-s^2)$. The value of noise is again independent of the signal level, and the assumption of equal spacing has resulted in simple independent expressions for both signal and noise.

It is now readily seen that the signal/noise ratio for a reference signal of the largest allowable amplitude when the pulses of the code group are all subject to a probability of error p is

$$\frac{P_{Sa}}{P_{Na}} = \frac{3}{8p}\frac{(1-2p)^2}{1-p}\frac{1}{1-s^2} \quad . \quad . \quad . \quad . \quad (26)$$

This may be multiplied by $f_r/2f_a$ to allow for the effects of the low-pass filter.

The signal power, noise power and signal/noise ratio are plotted on a decibel scale against p in Fig. 5, for the case when the number of digits in a code group is not too small ($s \ll 1$) and the audio band extends up to half the sampling repetition frequency. The zero level for signal and for noise is that of a reference tone of amplitude A in the absence of noise. On a linear scale of p the curves would be symmetrical about the value $p = \frac{1}{2}$, corresponding to completely random operation of the receiver. The signal/noise ratio is also shown in Fig. 6 (curve for one link), where the broken lines represent the approximation derived in Section 4. The error in this approximate formula amounts to 0·5 dB at a signal/noise ratio of 10 dB and to 0·05 dB at 20 dB.

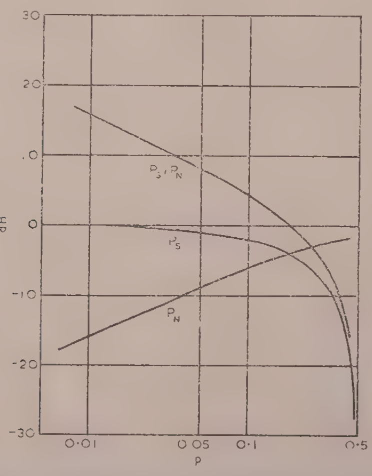


Fig. 5.—Variation of signal power P_S, noise power P_N, and signal/noise ratio P_S/P_N with probability of error p.

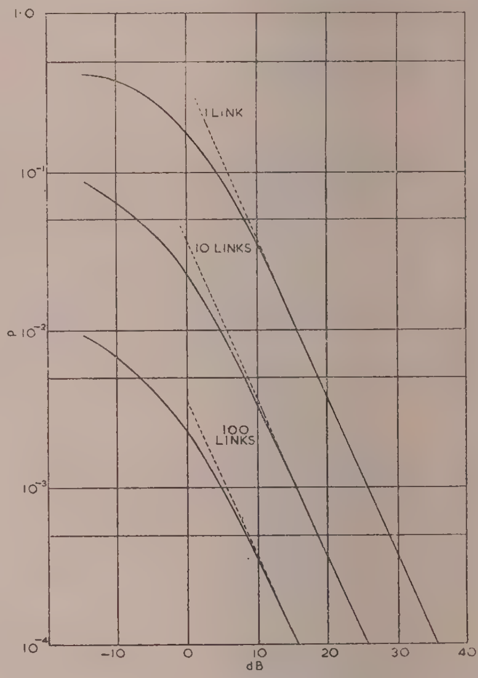


Fig. 6.—Variation of signal/noise ratio with p for systems of 1, 10, and 100 links.

(5.3) Multiple Links

If two or more links are used in tandem, the pulses being regenerated at each stage, there will be some reduction of the signal/noise ratio. If the probability p that a pulse is in error is small, after passage over M similar links the probability of an error is approximately Mp, the signal/noise ratio becoming

$$\frac{P_{Sa}}{P_{Na}} = \frac{3}{8Mp} \quad . \quad . \quad . \quad . \quad . \quad . \quad (27)$$

If p is not small, it is shown in Section 10.2 that the probability of a change after transmission over M similar links is

$$p_M = \tfrac{1}{2}[1-(1-2p)^M] \quad . \quad . \quad . \quad . \quad (28)$$

so that the accurate expression for the signal/noise ratio is

$$\frac{P_{Sa}}{P_{Na}} = \frac{3}{2}\frac{(1-2p)^{2M}}{1-(1-2p)^{2M}}\frac{1}{1-s^2} \quad . \quad . \quad . \quad (29)$$

where it must be remembered that p is the probability of an error after passage over only a single link. Curves of signal/noise ratio are plotted in Fig. 6 as functions of p for various numbers of links, assuming that $s \ll 1$ and $f_r = 2f_a$.

(6) SIGNAL/NOISE RATIOS AND THRESHOLD

After finding the signal/noise ratio in terms of p, the probability error, it is necessary to determine p in terms of the input ial/noise ratio. It is not difficult to do this for any practical tem, but it will be convenient here to assume the simplest sible transmission system. A double-current line-communiion system will be taken; see Fig. 7. The pulses, equally likely

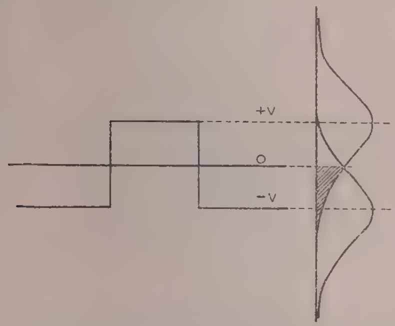


Fig. 7.—Pulse with superimposed probability distributions of Gaussian noise.

be positive or negative, will have amplitude V, but when subcted to a noisy transmission channel an amplitude V becomes bject to a random displacement which will be assumed to have Gaussian distribution, of r.m.s. value n_0, say.

The input signal/noise power ratio is readily seen to be

$$\left(\frac{P_S}{P_N}\right)_{in} = \frac{V^2}{n_0^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (30)$$

he probability of an error is given by the area passing the zero ie (shown shaded in Fig. 7), namely

$$p = \frac{1}{n_0\sqrt{(2\pi)}}\int_V^\infty \exp\left(-u^2/2n_0^2\right)du \quad . \quad . \quad . \quad (31)$$

he same probability of error is found if the pulse to be sent is gative.

Curves for this system are given in Fig. 8 for transmission over 10 and 100 links. They give the signal/noise ratio in the audio itput when the filter passes frequencies up to half the sampling

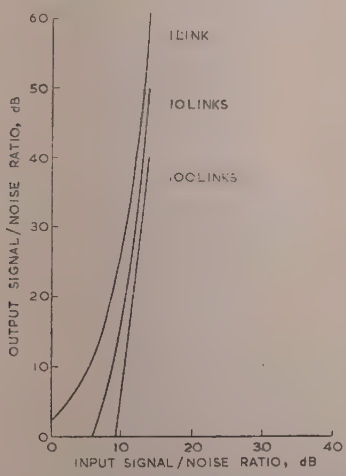


Fig. 8.—Variation of output signal/noise ratio with input signal/noise ratio.

frequency; for other filters the signal/noise ratio is to be increased by $10 \log (f_r/2f_a)$ dB. These curves differ markedly from those for amplitude modulation, frequency modulation, or the normal pulse systems. The main feature is the enormous slope of the curves such that in the working range an increase of input signal/noise ratio of 1 dB gives an improvement of 10–15 dB.

The output signal/noise ratio is low until the input signal/noise ratio exceeds about 10 dB, commonly known as the "threshold."

The curves for large numbers of links are shown to run quite closely alongside the curve for a single link. If they are examined closely, it will be found that at a given input signal/noise ratio the output value falls as the number of links is increased in the same way as if amplitude or frequency modulation were used. The advantage of pulse code modulation lies in the high slope of the curves, which means that the fall due to increasing the number of links may be offset by a surprisingly small increase in the power used for each link. This may be illustrated by taking the somewhat extreme case of 100 links in tandem. The signal/noise ratio corresponding to a single link may be restored, if pulse code modulation is used, by an increase in power of less than 2 dB in each link, whereas a more conventional system would need an increase of 20 dB.

(7) CONCLUSIONS

It has been shown that pulse code modulation has some decided advantages over conventional communication systems, particularly for transmission over a large number of links in tandem. The significant feature is the large increase of output signal/noise ratio consequent upon a change in input signal/noise ratio. It must not be forgotten, however, that inherent quantizing noise sets an upper limit to the signal/noise ratios attainable in speech transmission.

(8) ACKNOWLEDGMENT

Acknowledgment is made to the Chief Scientist, Ministry of Supply, and to the Controller of H.M. Stationery Office for permission to publish the paper.

(9) REFERENCES

(1) GOODALL, W. M.: "Telephony by Pulse Code Modulation," *Bell System Technical Journal*, 1947, **26**, p. 395.
(2) BENNETT, W. R.: "Spectra of Quantized Signals," *ibid.*, 1948, **27**, p. 446.
(3) OXFORD, A. J.: "Pulse-Code Modulation Systems," *Proceedings of the Institute of Radio Engineers*, 1953, **41**, p. 859.
(4) FITCH, E.: "The Spectrum of Modulated Pulses," *Journal I.E.E.*, 1947, **94**, Part IIIA, p. 556.
(5) AITKEN, A. C.: "Statistical Mathematics" (Oliver and Boyd).

(10) APPENDICES

(10.1) Evaluation of Output Signal and Noise

Let there be m independent events, each of which may be positive or negative. Let the probability that the rth event be positive be p_r. The m events represent the reception of the m pulses in a code group, and a positive event represents an error. Suppose that a positive event contributes an amount x_r and a negative one contributes y_r. Then the probability generating function of the amplitude of the resultant is

$$\text{p.g.f.} = \prod_{r=1}^{m}\left[p_r t^{x_r} + (1 - p_r)t^{y_r}\right]$$

The coefficient of t^z is the probability that the amplitude is z. The mean value and mean-square value of this amplitude are

required. Accordingly, if t is replaced by ε^u the moment generating function becomes

$$F(u) = \prod_{r=1}^{m} [p_r\varepsilon^{ux_r} + (1 - p_r)\varepsilon^{uy_r}]$$

The coefficient of $u^n/n!$ is the nth moment about zero. These coefficients are readily found by logarithmic differentiation:

$$\frac{F'(u)}{F(u)} = \sum_{r=1}^{m} \frac{x_r p_r \varepsilon^{ux_r} + y_r(1 - p_r)\varepsilon^{uy_r}}{p_r\varepsilon^{ux_r} + (1 - p_r)\varepsilon^{uy_r}}$$

$$\frac{F(u)F''(u) - F'^2(u)}{F^2(u)} = \sum_{r=1}^{m} \left\{ \frac{x_r^2 p_r \varepsilon^{ux_r} + y_r^2(1 - p_r)\varepsilon^{uy_r}}{p_r\varepsilon^{ux_r} + (1 - p_r)\varepsilon^{uy_r}} - \frac{[x_r p_r \varepsilon^{ux_r} + y_r(1 - p_r)\varepsilon^{uy_r}]^2}{[p_r\varepsilon^{ux_r} + (1 - p_r)\varepsilon^{uy_r}]^2} \right\}$$

Thus

$$F(0) = 1,$$

$$F'(0) = \sum_{r=1}^{m} [x_r p_r + y_r(1 - p_r)],$$

$$F''(0) - F'^2(0) = \sum_{r=1}^{m} [x_r^2 p_r + y_r^2(1 - p_r)] - \sum_{r=1}^{m} [x_r p_r + y_r(1 - p_r)]^2$$

$$= \sum_{r=1}^{m} p_r(1 - p_r)(x_r - y_r)^2$$

$F'(0)$ is the mean value and $F''(0) - F'^2(0)$ is the mean-square value about the mean.

To obtain the amplitude of the signal and the mean-square noise value in pulse code modulation write $p_r = p$. Then

$$\text{Mean amplitude} = p\sum_{r=1}^{m} x_r + (1 - p)\sum_{r=1}^{m} y_r$$

$$\text{Mean square about the mean} = p(1 - p)\sum_{r=1}^{m} (x_r - y_r)^2$$

But $\sum_{r=1}^{m} y_r$ is the undisturbed amplitude of the signal, a, and $\sum_{r=1}^{m} x_r = -a$ so that the mean amplitude in the presence of noise is $a(1 - 2p)$.

If $y_r = \pm A/2^r$, so that $x_r = \mp A/2^r$, the mean-square amplitude about the mean is

$$\overline{n^2} = 4p(1 - p)\sum_{r=1}^{m}\left(\mp\frac{A}{2^r}\right)^2$$

$$= 4p(1 - p)A^2\sum_{r=1}^{m}\frac{1}{2^{2r}}$$

$$= \tfrac{4}{3}p(1 - p)A^2(1 - s^2)$$

where $s = 1/2^m$.

(10.2) Transmission over a Number of Links

In a chain of M links, let the probability of the reversal of a code pulse in transmission over the mth link be p_m. Then form a probability generating function

$$G(t) = \prod_{m=1}^{M} [p_m t + (1 - p_m)]$$

such that the probability of w reversals after transmission over the chain is the coefficient of t^w. An odd number of reversals gives an error, while an even number gives no error. The overall probability p_M of an error is thus the sum of the coefficients of odd powers of t in $G(t)$, given by

$$p_M = \tfrac{1}{2}[G(1) - G(-1)]$$

$$= \tfrac{1}{2}\left[1 - \prod_{m=1}^{M}(1 - 2p_m)\right]$$

On substituting p_M for p in eqn. (26) the signal/noise power ratio after transmission over the chain is found to be

$$\frac{P_S}{P_N} = \frac{3}{2}\frac{\Pi(1 - 2p_m)^2}{1 - \Pi(1 - 2p_m)^2}\frac{1}{1 - s^2}$$

If the M links are similar, so that p_m has the same value p for each,

$$p_M = \tfrac{1}{2}[1 - (1 - 2p)^{2M}]$$

and

$$\frac{P_S}{P_N} = \frac{3}{2}\frac{(1 - 2p)^{2M}}{1 - (1 - 2p)^{2M}}\frac{1}{1 - s^2}$$

.396.677 : 629.135

Paper No. 1742 R
Mar. 1955

THE NOTCH AERIAL AND SOME APPLICATIONS TO AIRCRAFT RADIO INSTALLATIONS

By W. A. JOHNSON, M.A., Associate Member.

(*The paper was first received* 8*th April, and in revised form* 23*rd July*, 1954.)

SUMMARY

The results of a theoretical analysis, supported by experimental lasurements, show that very efficient aerials having interesting liation patterns and useful impedance characteristics can be obtained exciting metallic sheets with notches cut perpendicularly to an edge. he notch is about $\lambda/4$ deep, wide-bandwidth aerials can be obtained; t the detail of the polar diagram may be quite sensitive to frequency, pending on the shape and size of the sheet. If the notch is small mpared with the wavelength, it may still be tuned over a wide quency band and used as an efficient aerial. Examples are given the application of these small notches as aircraft aerials.

(1) INTRODUCTION

Considerable research and development work has been carried t by the Ministry of Supply on the design of aerials for high-eed aircraft,[1,3] and those discussed in the paper are based on e properties of an open-end slot, or notch, cut perpendicularly the edge of a metallic sheet. A brief mention of one application this type of aerial is made in a paper by Cary,[2] but the general operties have not previously been described. It is a very rsatile aerial, and can be used as a wide-band aerial when the mensions are similar to wide-band strip unipoles, or—more markably—as a very efficient way of exciting an aircraft as a diator even when the notch is very short in terms of the wave-ngth. The term "nitch" was coined in the laboratory con-rned with the development to distinguish these short aerials om notches of the order of a quarter of a wavelength in incipal dimension.

(2) DISCUSSION OF THE THEORY OF THE AERIAL

(2.1) The Theoretical Problem Considered

The performance of slot and dipole aerials may be related by e application of Babinet's principle, as first shown by Booker.[4] he Babinet analogue of a notch aerial is a unipole normal to the raight edge of the plane conducting sheet and lying in the plane the sheet. The problem chosen for theoretical analysis, by ason of its relative simplicity, was that of a current element ing in the plane of a semi-infinite perfectly-conducting sheet ormal to the edge.

Fig. 1 shows the positions of the current element and sheet lative to a system of spherical co-ordinates which are used to escribe the polar diagrams, and solutions were obtained for vo cases. The first gave the field anywhere, for a current ement at an infinitesimal distance from the sheet. The oblem was solved by verifying whether solutions which emed reasonable fitted the assumed boundary conditions. the second approach, expressions were derived for the istant field of a current element on the bisector of a corner flector, using the method of images in a corner of angle π/n, here n is an integer. If the angle is then made 360° ($n = \frac{1}{2}$), e current element is again perpendicular to the edge of a sheet,

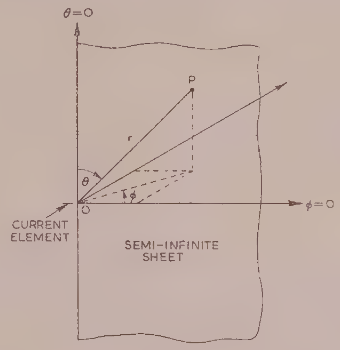


Fig. 1.—The co-ordinate system used to describe polar diagrams.

but this time the current element may be at a finite distance from the edge, and by superposition the correct distant field for any assumed current distribution on unipoles of finite size can be obtained. The transformation to the case where n is non-integral was later justified for the case $n = \frac{1}{2}$ by computing the distant field in the plane ϕ for $\theta = \pi/2$ in the co-ordinate system shown in Fig. 1, using a solution due to Sommerfeld for the diffraction pattern of a plane wave incident upon a semi-infinite sheet. The same polar pattern was obtained by both methods. Moullin[5] has met the same mathematical difficulty in his investigations on corner reflectors.

Computations based on this analysis have been made for aerials up to $0 \cdot 3\lambda$ long, the principal results of which are given in the following Sections. The practical examples given are restricted to the region where the aerial is short compared with a wavelength.

(2.2) Radiation Pattern of a Unipole on the Edge of a Semi-Infinite Sheet

For a current element at an infinitesimal distance from the edge

$$E_\theta = \frac{\cos\theta}{\sqrt{\sin\theta}}\sin\phi/2 \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$E_\phi = \frac{1}{\sqrt{\sin\theta}}\cos\phi/2 \quad . \quad . \quad . \quad . \quad . \quad (2)$$

These expressions use the co-ordinate system shown in Fig. 1, in which θ is measured in the plane of the sheet with the aerial element as origin, and ϕ is measured in the plane through the origin perpendicular to the sheet and its edge. E_θ and E_ϕ are the electric-field vectors parallel to these two planes.

It will be noticed that there is no radiation in the direction of the unipole on the side not occupied by the sheet ($\theta = \pi/2$). The field becomes infinite at the edge ($\theta = 0$) and is finite in the direction of the unipole on the same side of the sheet. These results may seem a little surprising, but are direct consequences of the assumption of a sheet of infinite extent.

Written contributions on papers published without being read at meetings are vited for consideration with a view to publication.
Mr. Johnson is at the Royal Aircraft Establishment.

(2.3) Radiation Resistance of a Unipole on a Semi-Infinite Sheet

For a unipole of finite length, l, carrying a sinusoidal distribution of current—which seems to be a fair approximation for unipoles up to $0 \cdot 3\lambda$ long—the radiation resistance referred to the base is

$$R_F = \tfrac{3}{8}\eta[C(\beta l) - \cos(\beta l)S(\beta l)]^2 \quad . \quad . \quad . \quad (3)$$

where η is the intrinsic resistance of free space as defined by Booker[4] and has the value 377 ohms, $\beta = 2\pi/\lambda$, and C and S denote the Fresnel integrals.

$$C_{(z)} = \int_0^z \frac{\cos t}{\sqrt{(2\pi t)}}dt \text{ and } S_{(z)} = \int_0^z \frac{\sin t}{\sqrt{(2\pi t)}}dt$$

If the unipole is so short that we can assume that the current distribution is linear, eqn. (3) reduces to

$$R_F = \frac{2}{3}\eta\frac{l}{\lambda} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

It is interesting to compare R_F with the radiation resistance of an ordinary unipole above a ground plane, referred to the base. Whereas the latter is proportional to the square of the length of the unipole for lengths short compared with a wavelength, in the present case R_F rises in direct proportion to the length. This leads to a reasonable value for the feed-point resistance, even for very short unipoles, and good efficiency in a practical aerial may therefore be reasonably expected.

(2.4) Reactance of a Unipole on a Semi-Infinite Sheet

With wire aerials it is usual to compute an approximate value for their input reactance by application of the theory of uniform or slightly non-uniform transmission lines. In both cases it is necessary to determine the average characteristic impedance of the aerial, and in the second case the deviation from the average is required. It has been calculated that the characteristic impedance Z_x at a distance x along the unipole from the sheet is $60 \log_\varepsilon(8x/d)$ and the average characteristic impedance, Z_{av}, $(\log_\varepsilon 8l/d - 1)$ for a cylindrical unipole of length l and diameter d.

For a rectangular strip of length l and width w the corresponding formulae are

$$Z_x = 60 \log_\varepsilon \frac{16x}{w} \quad \text{and} \quad 60 \log_\varepsilon \left(\frac{16x}{w} - 1\right)$$

(2.5) Current in the Sheet

The current has a radial component only. The current density becomes infinite at the edge of the sheet, the total radial current at any given distance from the origin remaining finite. The variation of current density with angle θ at a fixed distance from the origin is shown in Fig. 2.

(2.6) Notch Aerials on Semi-Infinite Sheets

The application of Babinet's principle for electromagnetic waves leads to the following results for a notch cut in the edge of a semi-infinite sheet. The radiation pattern is the same as that for the unipole, except for changes in the direction of polarization.

The radiation resistance is given by

$$R_F = \frac{2\eta}{3[C(\beta l) - \cot \beta_1 S(\beta l)]^2} \quad . \quad . \quad . \quad (5)$$

referred to the voltage V_F at the top of the notch. This relation

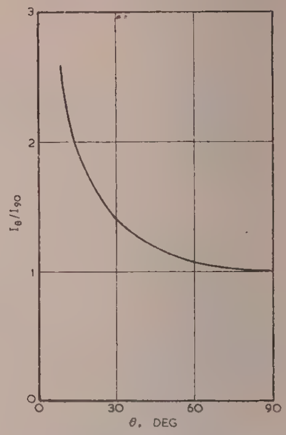


Fig. 2.—Variation of current density with angle at fixed distance from the origin.

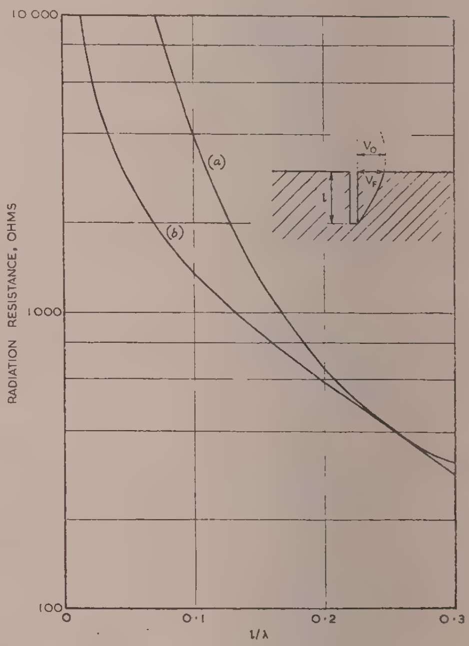


Fig. 3.—Radiation resistance of nitch aerials.

(a) R_0: referred to voltage loop.

(b) R_F: referred to top of nitch.

is plotted in Fig. 3, together with the radiation resistance R_0 referred to the voltage loop V_0. When the length l is smaller than $0 \cdot 1\lambda$ a good approximation is

$$R_F = \frac{3}{8}\eta\frac{\lambda}{l}$$

The reactance appearing in shunt across the mouth of the nitch may be taken approximately as that of a short-circuited transmission line having a characteristic impedance

$$Z_x = \frac{60\pi^2}{\log_\varepsilon \frac{16x}{w}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

an average characteristic impedance

$$Z_{av} = \frac{60\pi^2}{\log_e \frac{16l}{w} - 1} \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

...ere w is the width of the notch, l is its length, and x is the ...ance from the edge of the sheet at which Z_x is computed. ...e reactance is therefore easily computed.

...he line of current flow near the origin for a small notch—...itch—takes the simple form $C\sqrt[3]{\sin\theta}$, where C is a pro...rtionality factor. Two typical ones are shown in Fig. 4, and ...will be seen that the current has both a radial and tangential ...nponent.

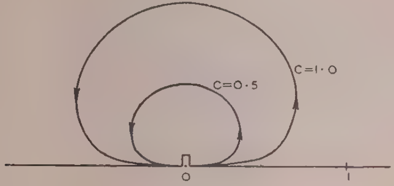


Fig. 4.—Typical lines of current flow in a sheet carrying a nitch.

(2.7) Notch Aerials on Finite Sheets

In practice, the sheet will always be finite in extent, and the ...bove theoretical results must be modified to suit a practical case ...efore it can be tested experimentally. For simplicity, only cases ...here the unipole or notch is symmetrical with respect to its ...xis will be considered.

By analogy with other aerial problems, e.g. a short unipole over ... finite counterpoise, it may be expected that the expression ...erived for radiation resistance is accurate within a few per cent ... the sheet extends for a distance of the order of a wavelength ...om the nitch. The reactance is due to the induction field, ...hich is significant only in the immediate vicinity of the notch ...r unipole and should therefore be independent of the size of ...e sheet if it is not extremely small.

The radiation pattern may be modified considerably on a finite ...eet, but in a sheet of three wavelengths extent it is assumed that ...e basic current distributions are unaffected near the aerial, and ...at reflections of current waves from the edges cannot contribute ...uch power to the radiation and will not change the basic nature ...f the diagram. Within the limit of these assumptions, some ...ualitative conclusions as to type of polar diagram can be drawn ...ithout detailed computation.

The diagrams to be expected are shown for the three principal ...ections in Fig. 5.

For the semi-infinite sheet there is no component E_θ on the ...ide of the sheet, and E_θ has opposite directions on the two sides. Therefore, there should be no field immediately to the rear of the ...otch. There will be some diffraction in the rear sector. The ...eros for E_ϕ in the direction of the edge may be explained by ...he cancellation of the fields on the two sides of the sheet, or by ...onsidering the current components parallel to the edge.

Because the reflected waves have been neglected, there will be more or less pronounced fluctuations on these basic patterns, depending on the actual dimensions of the sheet. One particular case has been computed for one cross-section of the radiation pattern for one direction of the polarization by the application of first-order diffraction theory. The example chosen was that of a semi-circular sheet of radius $1 \cdot 75\lambda$ with the notch in the centre of the straight side; the result of the computation is given in Fig. 6.

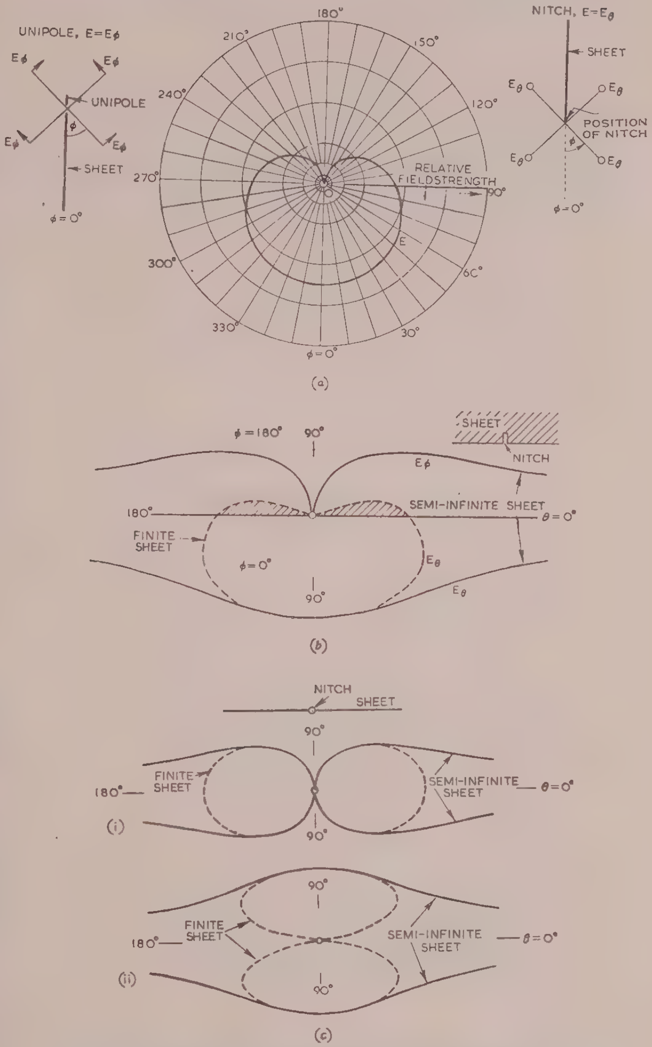


Fig. 5.—Polar diagrams of a nitch aerial.

(a) In plane $\theta = \pi/2, -\pi/2$.
(b) In plane $\phi = 180°, 0°$.
(c) In plane $\theta = 0°$.
(i) E_ϕ.
(ii) E_θ.

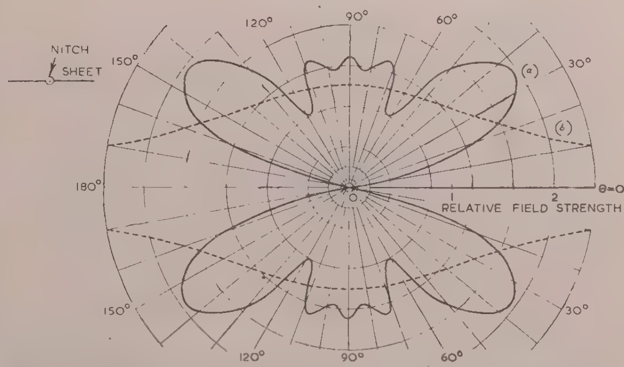


Fig. 6.—Theoretical polar diagram for a nitch in a semicircular sheet.

(a) Sheet of radius $1 \cdot 75\lambda$.
(b) Semi-infinite sheet.

(3) EXPERIMENTAL RESULTS ON UNIPOLE AND NOTCH AERIALS

(3.1) Impedance Measurements

A unipole was mounted above the top edge of a vertical sheet of tin-plate 9 ft square, and impedance measurements were made. The sheet could be extended to twice its horizontal extent by adding extra pieces, but no significant change of impedance could be measured, so confirming the assumption made in the theoretical discussion. A typical set of measurements is given in Fig. 7

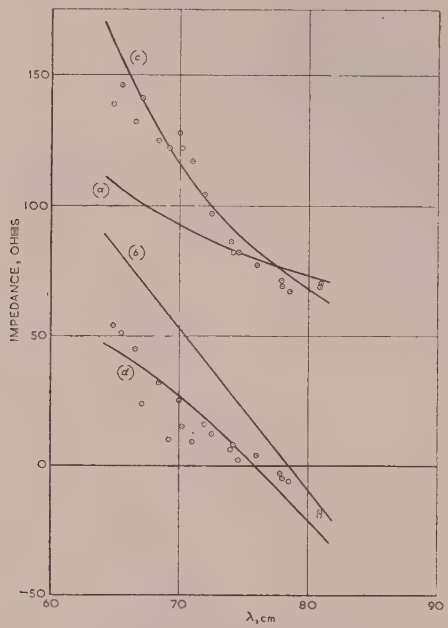


Fig. 7.—Impedance of unipole on the edge of a sheet.

(a) Computed resistance without shunt capacitance.
(b) Computed reactance without shunt capacitance.
(c) Computed resistance with 1 μμF shunt capacitance.
(d) Computed reactance with 1 μμF shunt capacitance.
⊙ ⊙ ⊙ Measured values.

for a unipole 2 mm in diameter and 18·2 cm long in the frequency band 400–500 Mc/s. Curve (*a*) is the theoretical radiation resistance from eqn. (3) and curve (*b*) is the reactance obtained from the slope

$$\frac{\partial x}{\partial \lambda} = -Z_{av}\frac{2\pi l}{\lambda_0^2}$$

of the reactance curve in the vicinity of the first resonant wavelength, assuming a "shortening" effect on the unipole of 8%. Z_{av} is found from the theoretical expression already given to be 335 ohms. The measured values are seen to be of the correct order, but do not follow the theoretical curves (*a*) and (*b*) very closely. The theory neglects unavoidable base capacitance in the join between the unipole and the measuring cable. If it is assumed that there is an effective base capacitance of 1 μμF and the theoretical values are recalculated, curves (*c*) and (*d*) are obtained. These are a reasonably good fit to the measurements.

(3.2) Radiation-Pattern Measurements

Measurements were made of the three main cross-sections of the polar diagram for a notch and a unipole on a rectangular sheet measuring 8 ft 11 in by 3 ft at 390 Mc/s, corresponding in size with the assumptions made in the theoretical discussion. The notch was 5 in long and $\frac{3}{4}$ in wide, and was fed by a concentric cable. The inner conductor of the cable was connected to one side of the notch and the outer to the other side. The unipole was 6 in long and about $\frac{1}{8}$ in diameter, and was fed directly from a cable. In order to prevent radiation from currents on the outer sheath of the concentric cable, it was decoupled by a quarter-wave choke where it left the sheet. The patterns were in good general agreement with theoretical predictions, typical examples being shown in Fig. 8.

The polar diagram for the notch in a rectangular sheet was measured at various frequencies, and while the general main characteristics remained similar to that predicted in the approximate theory given in Section 2.7, the variation in detail was considerable, and varied rapidly with frequency, particularly in the planes normal to the notch axis. Measurements on a semi-circular sheet were carried out, and here again the detail varied rapidly with frequency. The pattern most similar to the computed one (Fig. 6) is shown in Fig. 9, which was obtained at a frequency 2% lower than predicted. As the theory is a first-order approximation, this may be acceptable as good agreement. Another interesting point is that very similar diagrams may be obtained from semi-circular and rectangular sheets if the diameter is about 5% longer than the side of the rectangle.

(4) PRACTICAL APPLICATIONS OF THE SHORT NOTCH, OR NITCH, AERIAL TO AIRCRAFT

(4.1) Vertical Polarized System at Very High Frequencies

The nitch aerial is essentially a device with a fairly high Q-factor, and therefore in its practical applications may be used either as a fixed tuned aerial over a narrow frequency band, or in association with a tuning unit over a wider frequency band.

For single-frequency reception in the region of 100 Mc/s a very small nitch, only about 3 in long, is adequate. Since many v.h.f. systems operate with vertical polarization, it is natural to try to excite the vertical tail-fin of an aircraft as an aerial.

For purposes of test a full-size tail-fin for a small aircraft, 34 in high and 24 in wide, was made. A 7 ft square of sheet metal mounted upon a wooden frame was used as earth plane, with the fin mounted at the centre. The whole could be rotated so that polar diagrams could be measured.

A comprehensive model of a complete aircraft was not attempted in this experiment, as while modifications to polar diagram shapes by aircraft structure are inevitable, it was more important to establish the basic polar diagram of the fin radiator than to study aircraft effects. Two nitches, $3\frac{1}{2}$ in deep and 1 in wide, were cut in the leading edge of the fin, spaced 4 in apart, with the lower one 4 in from the counterpoise.

The nitches were tuned and fed according to the circuit shown in Fig. 10. The construction adopted was a copper shell which fitted the hole cut into the fin, supported by a solid block of Perspex machined to the fin contour. This makes a compact aerial unit which is easily manufactured, transported and assembled. In a v.h.f. nitch about $\lambda/30$ long a filling of Perspex has no measurable electrical effect.

The reactance of a nitch of characteristic impedance Z_0 and electrical length θ is $Z = jZ_0 \tan\theta$, which is approximately $jZ_0\theta$, since θ is small. When the nitch is filled with material of permittivity ϵ, the characteristic impedance is reduced by $\sqrt{\epsilon}$, and the electrical length is increased by $\sqrt{\epsilon}$. Therefore

$$Z \simeq \frac{Z_0}{j\sqrt{\epsilon}}\sqrt{(\epsilon)}\theta$$

i.e. the reactance is unchanged by a solid dielectric filling. The dielectric loss in the Perspex was insignificant at these frequencies.

Apart from the convenience of having a solid unit to handle,

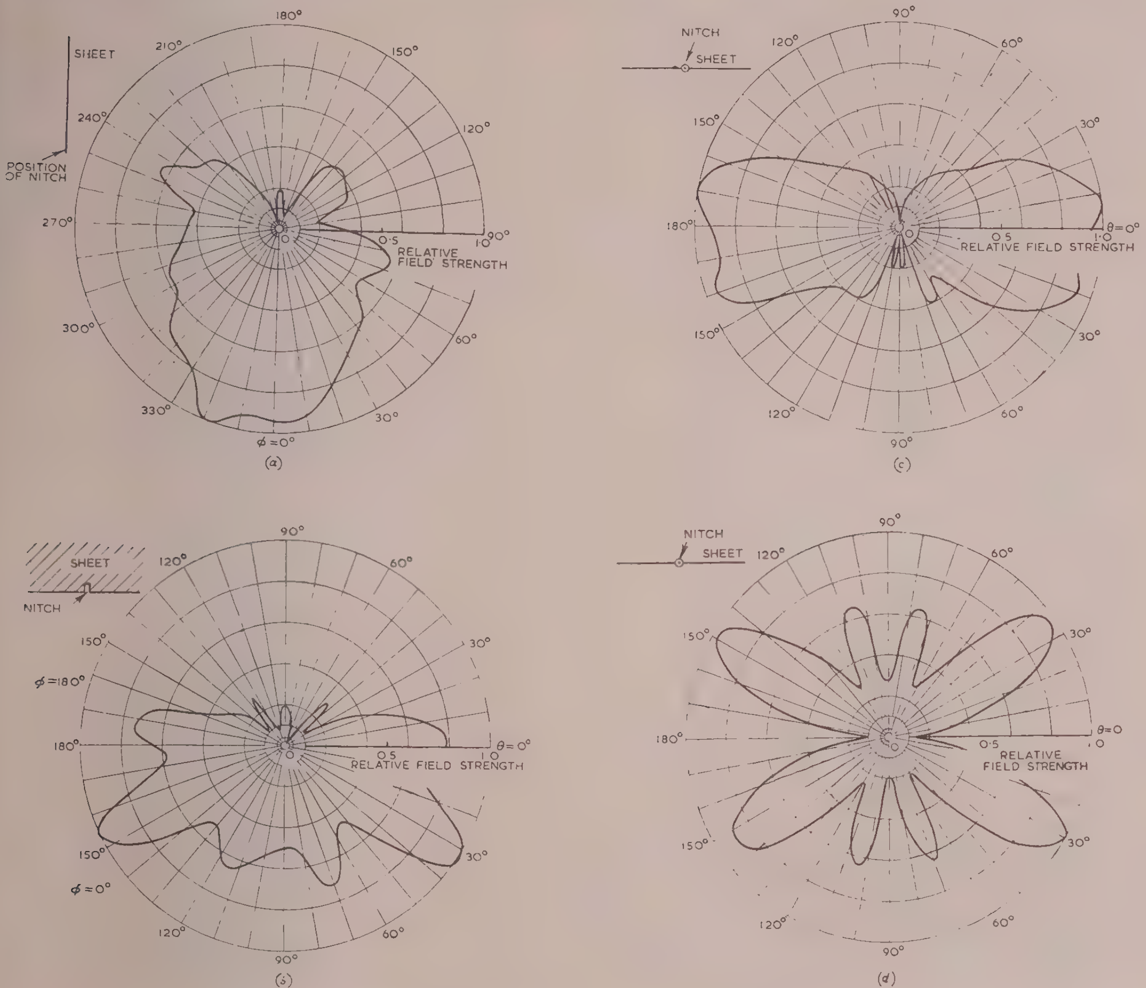


Fig. 8.—Measured polar diagrams of a nitch aerial in a rectangular sheet.

(a) Normal to sheet.
(b) In plane of sheet.
(c) E_ϕ normal to axis of nitch.
(d) E_θ normal to axis of nitch.

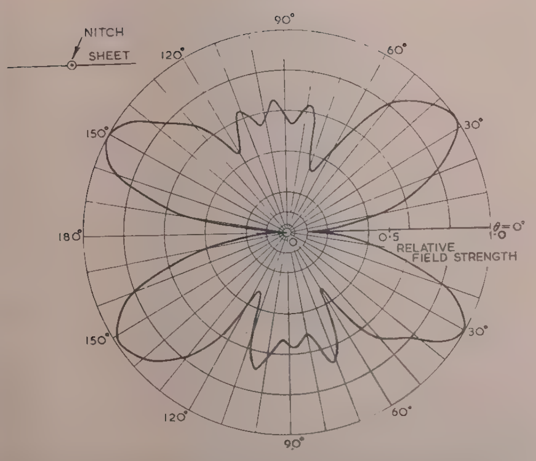


Fig. 9.—Measured polar diagram of a nitch aerial in a semicircular sheet.

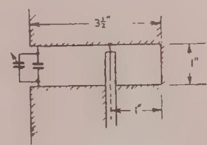


Fig. 10.—Dimensions and circuit of a nitch aerial.

a solid filling provides rigid support for the tuning condenser and some stiffening of the leading edge and skin of the fin.

The radiation resistance of the nitch when mounted in the fin was about 1 000 ohms and the Q-factor was about 35. This is of the same order as predicted by theory for a nitch in a lamina, and feed points were easily found for a 50-ohm cable. Although the two nitches were only 4 in apart, provided that they were not tuned to the same frequency there was no mutual interference. In one experiment one nitch was used for transmission while the other was used simultaneously for reception in a 2-frequency duplex system. There was less transmitter-receiver interference with this arrangement than there was when using the same

equipment on two whip aerials 6ft apart. This, of course, is due to the high Q-factor of the nitches, but the experiment indicates one way of using the aerial's properties to good effect.

The polar diagram of the nitch in the experimental rig is given

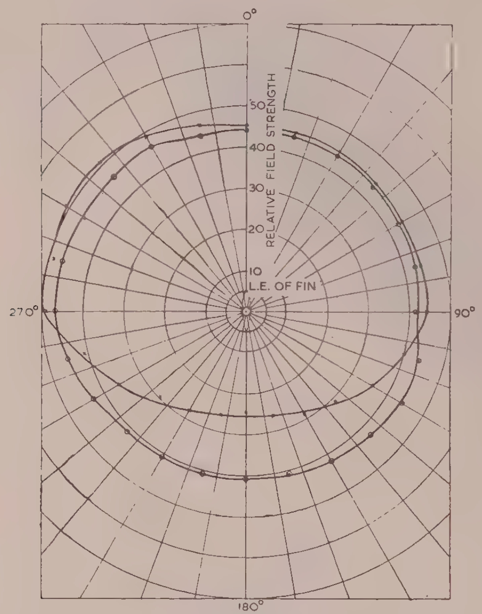


Fig. 11.—Comparative polar diagrams of nitch and whip aerials.

—×—×— Nitch aerial.
—○—○— Whip aerial.

in Fig. 11, together with that of a whip aerial on the same earth plane. Integration of the polar diagram shows that the mean gain of the nitch in the horizontal plane is 88% of that of the whip. The polar diagram of the nitch is sufficiently close to a circle to be perfectly acceptable for aircraft communication purposes, as subsequent trials of the unit in aircraft have shown.

(4.2) A Horizontally-Polarized System

One of the internationally agreed navigational aids for aircraft is a v.h.f. instrument landing system which requires the aircraft to be able to receive a horizontally-polarized signal in the band 112–118 Mc/s with substantially uniform sensitivity in all azimuthal directions. If the aerial is sensitive to vertically polarized radiation, errors due to spurious vertically-polarized radiation from the ground aerial may become apparent. These signals mark out an approach path for the aircraft in azimuth. Similar requirements are called for at 400 Mc/s for an approach path in elevation, known as the glide path.

It is therefore most important that the aircraft aerials, particularly in the forward sector in approach attitude, should conform very closely to ideal horizontally-polarized aerials.

A notch in a wing has excellent polarization characteristics, and provided that it can be mounted to give a reasonably circular polar diagram, provides an excellent solution to the problem.

The position should be on, or as close as possible to, the wing tip of the aircraft, and typical polar diagrams are given in Fig. 12. It will be noticed that when the notch is actually at the tip the polar diagram is, for practical purposes, reasonably circular, but in the leading-edge position a much less circular pattern is produced. For positions further inboard still, a much more pronounced lobe structure, varying rapidly with frequency and quite useless for omnidirectional applications, is produced.

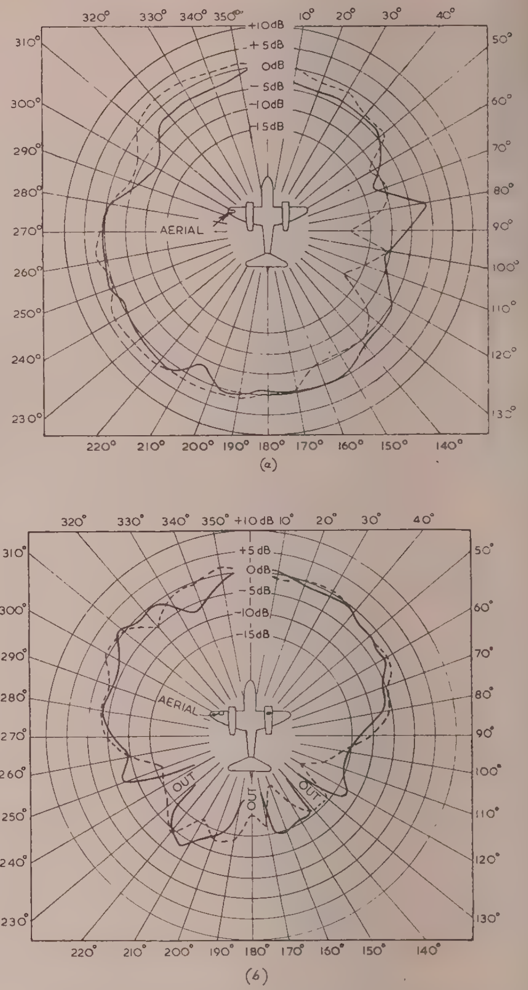


Fig. 12.—Polar diagrams for nitch aerials in a small aircraft.

———— In the air.
— — — — On the ground.
(*a*) Wing-tip aerial.
(*b*) Leading-edge aerial.

These diagrams were taken on a full-scale aircraft in the course of the engineering development of the aerial, and illustrate the degree of agreement which is normally acceptable for practical purposes between ground and air tests.

Provided that the ground measurements are made on a large enough site with reasonable precautions, sufficient information can be obtained to predict the polar diagram obtainable in the air to this order of accuracy from a simulated installation of the aerial. This is a most important practical consideration when one bears in mind the high cost of completing a flyable installation in an aircraft and carrying out the flight tests.

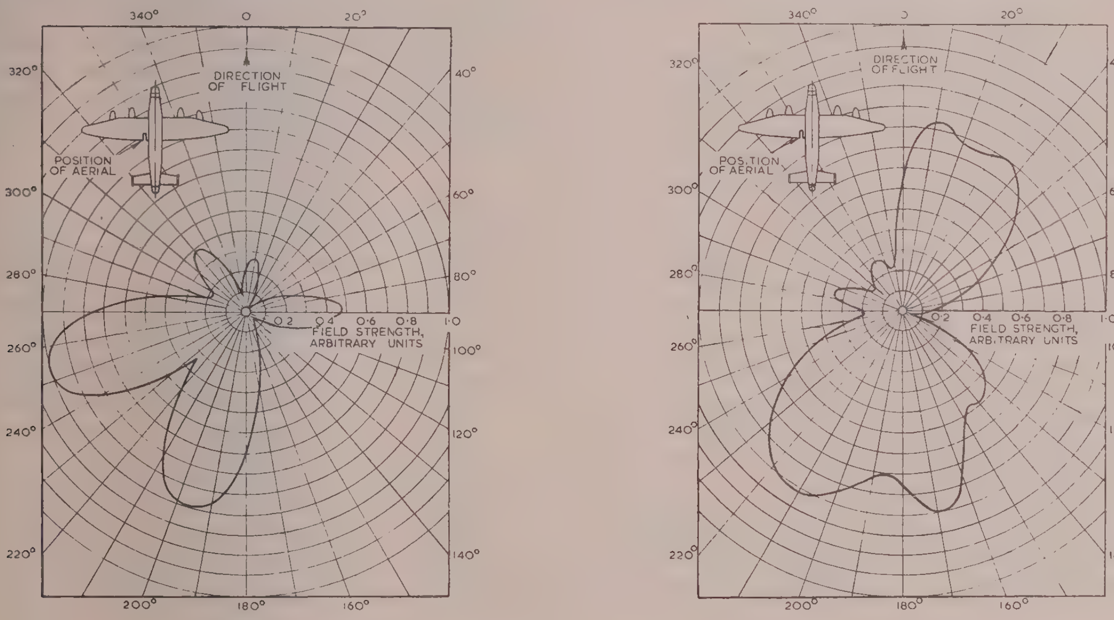
Having established the principle that a notch in the wing tip would produce an acceptable polar diagram, it was necessary to expend a considerable amount of engineering effort in applying the basic principles of the notch aerial to a design which was at once easy to mass-produce, applicable to a wide variety of wing sections, and of minimum size, to meet the impedance/bandwidth requirements. The solution obtained was to make a U-shaped framework of steel, tuned with a preset condenser across the mouth, and fed by a tap terminating in a socket

which a connector could be plugged after installation. A ond condenser was inserted in series with the tap to improve impedance match over the frequency band. The first-order ct of mounting this unit in wings of different cross-section is vary the characteristic impedance of the notch, but not the iation resistance. Consequently, a selection of the correct ue of the preset condenser across the mouth is sufficient to apt a universal unit to many aircraft. The unit is installed the wing tip near the wing-tip light, and the notch is afterrds covered with a piece of glass-fibre laminate screwed into edge, thus restoring the contour of the wing and adding to stiffness of the structure. A similar design, but of only onearter the size, has been equally successful for the 400-Mc/s uirement for the glide-path application.

(4.3) The Notch Aerial for H.F. Communication Systems

The notch aerial can make a contribution to the problem of providing suppressed aerials for aircraft transmission in the h.f. communication band from about 2·5 to 20 Mc/s. If the aircraft has a large tail-fin it is possible to excite this by inserting a notch somewhere near the base and so obtain a polar diagram very similar to those described by Cary[3] for systems using insulated tails. The notch has the great structural advantage that no reliance is placed on a strip of insulating material to carry structural loads, and no elaborate precautions against structural damage by lightning need be taken, since the whole tail structure is one connected piece of metal.

Less satisfactory polar diagrams, but none the less quite useful ones at many frequencies in the band, can be obtained by inserting



(a) (b)

Fig. 13.—Polar diagrams for nitch aerials in a Lancaster.

(a) At 15·055 Mc/s.
(b) At 7·308 Mc/s.

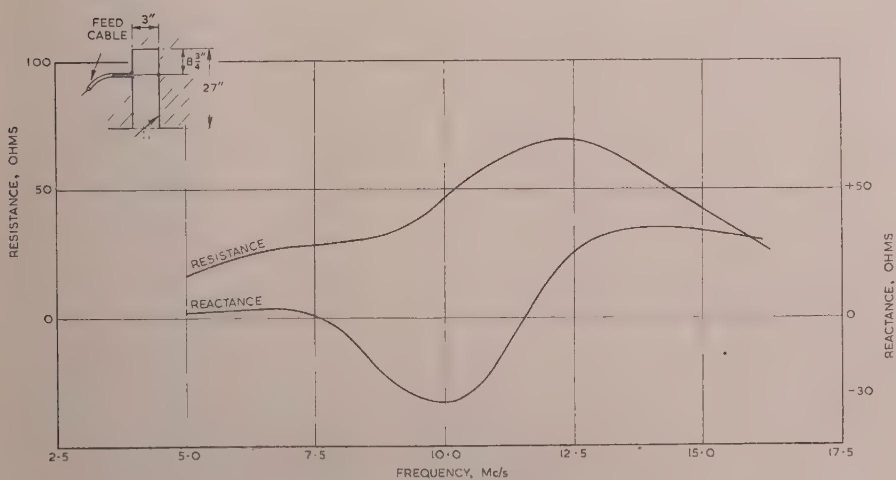


Fig. 14.—Input impedance of h.f. notch.

a notch in the trailing edge of the wing, where the aerodynamic loads are light. Here the radiation is predominantly horizontally polarized, which is of no practical consequence when long-distance transmission involving ionospheric reflection is involved. Experience so far obtained with excited wing aerials shows that there is normally sufficient vertically-polarized radiation to give usable direct-ray ranges in the majority of cases with vertical ground aerials. Extensive flight trials have also shown that although the polar diagram in the azimuthal plane is very irregular, at long ranges communication may be maintained quite surprisingly well for all aircraft azimuths.

The conclusion drawn, therefore, is that, while wing excited aerials are not by any means ideal, they are not as unreliable as might be thought from an examination of the polar diagram in the azimuthal plane for horizontally-polarized radiation. A typical pattern produced by an experimental notch installation in the trailing edge of a Lancaster aircraft is given in Fig. 13(*a*). In most cases the patterns are quite asymmetric and irregular. It was only by very careful selection of frequency that a pattern approximating to that of the wing radiating as a horizontal dipole could be obtained, and such a special case is shown in Fig. 13(*b*).

For structural reasons, the greatest depth of nitch which aircraft designers are normally prepared to accept with reasonable equanimity is about 40 in, and a typical curve of input resistance and reactance of such a deep nitch for a fixed feed-point is given in Fig. 14. The resistance falls quite rapidly at frequencies lower than 4 Mc/s; the lower limit at which it is possible to use an aerial of these dimensions depends upon the losses in the matching unit, the efficiency acceptable, the particular configuration of the aircraft and, for the transmitting aerial, the voltage which can be accepted across the aerial system when tuned to resonance.

For frequencies above 2·5 Mc/s successful tuning units have been evolved. The principle of feeding the nitch is the same as at higher frequencies, but now, as the frequency range is 2·5 to 20 Mc/s and the Q-factor of the nitch is high, it is necessary to retune the nitch at each frequency and provide for this to be done either automatically in step with frequency-change mechanism in the transmitter or manually by an operator. It is necessary to enclose the tuning unit in a container filled either with oil or air at atmospheric pressure or higher to prevent corona discharges at high altitudes. The unit is tucked inside the aerodynamic fairing directly beside the nitch which it tunes, and the nitch itself would be quite invisible when faired in with insulating material and given a surface finish. These and other applications of the nitch aerial are protected by patent.[6]

(5) CONCLUSION AND ACKNOWLEDGMENTS

The investigations carried out as described in the paper for exciting aircraft wings and similar structures have provided alternative and improved solutions to some of the problems of designing suppressed aerials for modern aircraft. The method is versatile and has been already applied successfully to systems as diverse in frequency as 2·5 Mc/s and 500 Mc/s.

Such an extensive programme as the one described has received assistance from many sources inside the Royal Aircraft Establishment, and a general acknowledgment is therefore made to the staff concerned. In particular, the author is indebted to Mr. P. Rothe for very extensive analytical and computational work on the theoretical examples. The paper is published by permission of the Chief Scientist, Ministry of Supply, and the Controller, H.M. Stationery Office.

(6) REFERENCES

(1) Johnson, W. A.: "Recent Developments of Aircraft Communication Aerials," *Journal I.E.E.*, 1947, **94**, Part IIIA, p. 452.
(2) Cary, R. H. J.: "The Slot Aerial and its Application to Aircraft," *Proceedings I.E.E.*, Paper No. 1136R, July, 1951 (**99**, Part III, p. 187).
(3) Cary, R. H. J.: "A Survey of External and Suppressed Aircraft Aerials for use in the High-Frequency Band," *Proceedings I.E.E.*, Paper No. 1173R, September, 1951 (**99**, Part III, p. 197).
(4) Booker, H. G.: "Slot Aerials and their Relation to Complementary Wire Aerials (Babinet's Principle)," *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 620.
(5) Moullin, E. B.: "Theory and Performance of Corner Reflectors for Aerials," *ibid.*, 1945, **92**, Part III, p. 58.
(6) British Patent No. 704659, 1954.

.016 : 537.311.62 : 531.717.8

Paper No. 1756R
Mar. 1955

A SURFACE-TEXTURE COMPARATOR FOR MICROWAVE STRUCTURES

By A. F. HARVEY, D.Phil., B.Sc.(Eng.), Member.

(*The paper was first received* 14*th April, and in revised form* 21*st August*, 1954.)

SUMMARY

n account is given of the influence of surface texture on the per- nance of microwave components and instruments. Methods of isurement are discussed, and a description of a simple scale for nparing surface textures is given, together with brief details of nufacture. The range of surface finishes and the methods of rication cover those used in modern manufacturing methods in this d.

(1) MICROWAVE SURFACES

n microwave structures[1] the most important parts are the lumes in or around which the electric and magnetic fields st and the surfaces carrying the radio-frequency currents. At se high frequencies the current amplitude decays very rapidly th distance from the surface towards the interior of the con- ctor, and for practical purposes the entire current can be nsidered to flow in an effective thickness close to the surface med the skin depth. This thickness, which is inversely pro- rtional to the square root of the frequency, is given in Table 1 r high-conductivity copper, the units being microinches. For mparison, the wavelength of green light is 22 microinches, hile at 50 c/s the skin depth is approximately 0·5 in.

Table 1

Skin Depth for Copper at Various Frequencies

Frequency, Gc/s	3	10	30	100	300
Wavelength, mm	100	33	10	3·3	1·0
Skin depth, microinches	50	29	16	9	5

The thickness is also proportional to the square root of the sistivity of the conducting metal, and for other materials must multiplied by the factor given in Table 2. It has been shown[2]

Table 2

esistivity and Multiplier for Skin Depth for Various Metals

Material	Resistivity, microhm-cm	Multiplier
Silver	1·56	0·98
Copper	1·62	1·00
Silver, 7½% Copper	1·80	1·06
Aluminium	2·83	1·30
Brass, 90% Copper	3·90	1·55
Brass, 70% Copper	6·52	2·00

at the attenuation coefficient of microwave structures is versely proportional to the skin depth. This ignores changes attenuation due to imperfections in the surface caused by ulty finishing or treatment[3,4] and those occurring in special conditions, such as at very low temperatures, when the mean free path of the conducting electrons is comparable with the skin depth.[5]

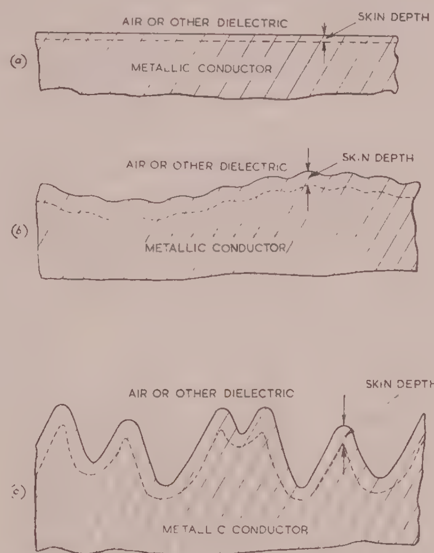


Fig. 1.—Increase of current path-length with surface roughness.

(*a*) Surface roughness zero.
(*b*) Surface roughness small compared with the skin depth. The long period secondary texture is also shown.
(*c*) Surface roughness large compared with the skin depth.

It will be seen from Fig. 1 that unless the surface is smooth to the order of the skin depth the currents will traverse longer paths. Thus the attenuation of a wave travelling along the line will be increased, or the Q-factor of a resonant structure will be decreased. It has, for example, been shown by S. P. Morgan[6] that for a surface roughness equal to the skin depth the attenuation coefficient is increased by a factor 1·60; when it is twice the skin depth the factor is 1·80, and for one-half the skin depth it is 1·20. Thus in all cases where low attenuation is of importance the surface roughness should be kept at a minimum and preferably not exceed a value equal to about one-half the skin depth.

In the past, it has been customary in reports, drawings and specifications relating to microwave structures to describe the current-carrying surfaces in such terms as smooth, mirror finish, highest possible finish, etc. These terms have proved unsatisfactory and this fact has been one of the reasons why the manufacture of microwave structures cannot be expressed in ordinary engineering terms. Recently, however, there has been a tendency[7,8,9] to specify the surface texture in microinch limits. This is mainly due to the more severe performance requirements of microwave equipment and the extension of the working wavebands to the millimetre wavelength region. Thus it is evident that manufacturers and others concerned with microwave structures must have some means of measuring or comparing surface finishes of this order.

Written contributions on papers published without being read at meetings are vited for consideration with a view to publication.
Dr. Harvey is at the Radar Research Establishment.

The increase in the current path-lengths is mostly due to the primary texture, i.e. the roughness which results from the action of the tool. The effect of secondary texture, such as long-period waviness, resulting from imperfection in the performance of the tool, is small, provided that, as is usual, the depths of the irregularities are much less than a free-space wavelength and are within the dimensional tolerances required.

An ideal machine would measure the increase in the linear dimension of the surface caused by the roughness, but the author is not aware of such a device. The normal practice[10,11] in this country is to traverse a mechanical stylus, the radius of whose tip does not exceed 0·0001 in, over a short sampling length of the surface, the deviations of the stylus, after amplification, being recorded on a moving chart or displayed on an integrating meter. Such a meter gives what is termed the centre-line-average height of the irregularities. An alternative basis for assessment is the r.m.s. value of the surface roughness, which is more commonly used in the United States. For most types of surface this is $\pi/2\sqrt{2}$ or 1·11 times[11] the centre-line-average figure, a factor which, in practice, can be ignored. Very fine surface textures can also be determined by optical interference; the lack of parallelism of fringes between the surface under test and that of an optically flat glass plate is a measure of the roughness.

Such machines as described above are expensive and are usually installed in laboratories or standard rooms remote from the workshop. A much more convenient but not, of course, such an accurate method is to employ comparator plates of known surface finish. These have been used in the United States for general engineering work, but the finishes are too coarse and are not of the right type for microwave work. The grades of surface texture have recently been standardized,[11] the centre-line-average figures ranging from 1 to 1 000 microinches. For microwave work, values ranging from 2 to 63 microinches are usual and have been chosen, with various methods of finishing, for the comparator scale described below.

(2) SURFACE-TEXTURE COMPARATOR SCALE

The shapes of microwave structures are complicated and a variety of operations are used in their manufacture. Some examples, with approximate surface finishes, are given in Table 3.

In those operations which produce a surface texture having a directional quality or lay the method of manufacture should be such that this is parallel to the current flow where this is unidirectional. In other cases, the method of finishing or fabrication will depend on the shape of the structure and on the dimensional accuracy. The latter requirement almost invariably excludes any polishing or buffing operation, although these can give very smooth surfaces. Such operations are, in any case, not recommended for microwave structures since they tend to form an amorphous low-conductivity Beilby layer,[12] which may have a thickness up to one microinch.

Most of the listed methods of fabrication can be put into one of three classes. The first class has little or no lay and includes grinding, lapping, honing, super-finishing, polishing as well as casting methods and powder metallurgy. The second is carried out with sharp pointed tools and has a lay in the direction of the movement of the work with respect to the tool, irrespective of which is fixed and whether the motion is in a straight line or in an arc. Operations in this class include turning, boring, fly-cutting, planing; similar textures are given by extrusion, drawing and hobbing. In the third class the tool is broad and the lay tends to be parallel to the width of the tool and at right angles to the relative movement of the work. This class includes roller milling, broaching, reaming, scraping and burnishing.

Table 3

Surface Finish for Various Methods of Fabrication

Method of machining or fabrication	Microinches
Turning, preliminary finish	63 to 125
Turning, ordinary finish	32 to 63
Turning, fine, ferrous metals ..	8 to 32
Turning, fine, non-ferrous metals ..	4 to 16
Diamond turning, non-ferrous metals	2 to 8
Boring, ordinary	16 to 32
Boring, fine, ferrous metals	8 to 16
Boring, fine, non-ferrous metals ..	4 to 8
Diamond boring, non-ferrous metals	2 to 4
Planing, ordinary	16 to 63
Planing, fine	8 to 16
Extrusion or drawing, mirror finish..	8 to 32
Hobbing	8 to 32
Milling, ordinary	32 to 63
Milling, fine	8 to 32
Reaming, ordinary	16 to 32
Reaming, fine	4 to 16
Broaching, ordinary	16 to 32
Broaching, fine	4 to 16
Scraping	8 to 32
Burnishing	2 to 8
Grinding, ordinary	16 to 32
Grinding, fine	8 to 16
Grinding, superfine	2 to 8
Honing	1 to 8
Lapping	1 to 4
Super finishing..	1 to 4
Polishing	1 to 2
Casting, investment	63 to 125
Casting, polished dies	4 to 32
Powder metallurgy	63 to 125

In order to facilitate comparison of surface textures, it is desirable that the comparator scale should include all three different classes of finish, since the values given refer to measurements in a direction approximately at right angles to the lay. This gives a different appearance and sense of touch to the surfaces.

After much consideration and the preparation of samples, the following series were chosen as being the most useful in microwave work, the finishes being given in microinches (centre-line-average): lap 2, grind 4, 8, 16, 32 and 63, turn 8, 16, 32 and 63, mill 32 and 63. The comparison was intended to be done by sight and touch, and since most microwave surfaces are metallic it was further decided that the comparator scales should also be in metal rather than in materials such as plastics.

The individual production of master scales, whilst probably the best method, would have made them so expensive that only a very few could have been made. It was therefore decided to produce them as copies by electro-deposition from a master, the simplest method being the technique followed in the production of gramophone records. Thus the appropriate finishes were put on the ends of cylindrical plugs, which were then mounted in a circular copper disc. In order to obtain the highest finishes, and to prevent errors due to corrosion, these plugs were made of stainless steel. To reduce the cost of production three identical sets of finishes were mounted on the disc, and after this had been engraved with lettering and border lines the final appearance was as in Fig. 2.

From this master disc a negative copy was first of all made by electro-deposition in nickel, the reverse side being masked off. From this negative, pressings could have been obtained in plastic, but, in the method adopted, positive copies were made in copper by further electro-deposition. For permanence, the face of the copies were given a chromium flash and to give rigidity the plates

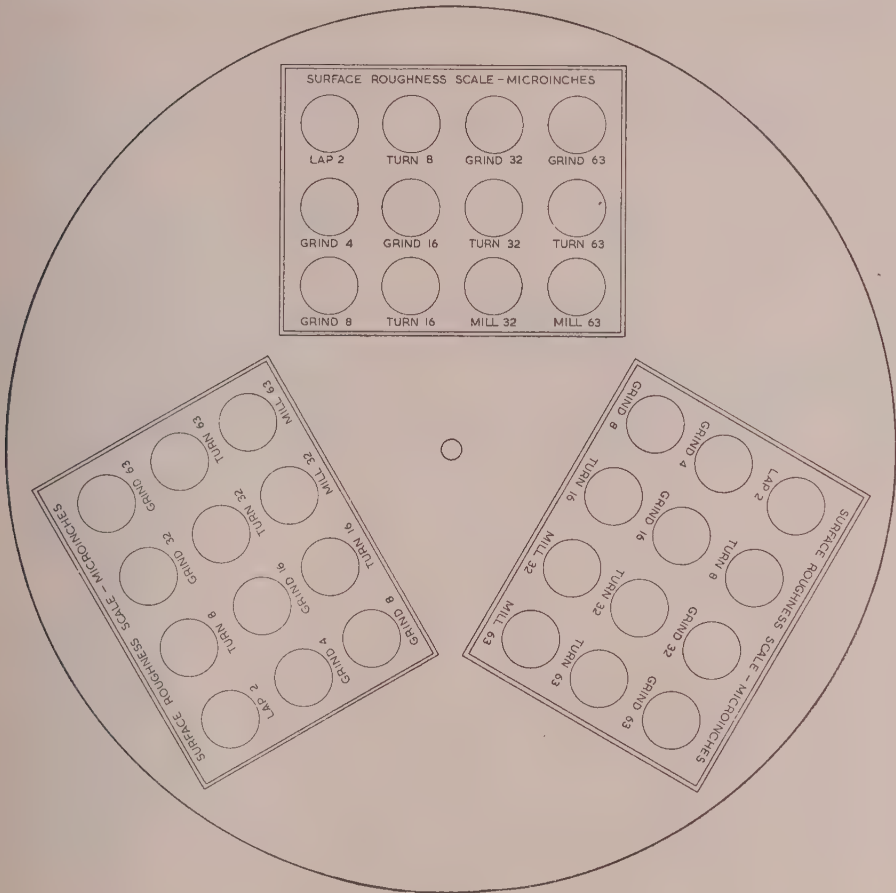


Fig. 2.—Copper disc with master roughness scales.

The diameter of the disc is 12in with thickness $\frac{1}{8}$in. Three identical scales are incorporated for economy in production. The roughness patches are on stainless-steel inserts.

were backed by soldering to them thin copper discs to give a total thickness of about $\frac{1}{16}$in. The scales were then cut out and trimmed to the outer border line to form the completed article shown in Fig. 3.

These scales have proved convenient to use and some mention might be made of the accuracy of measurement. The main error is in the uncertainty of comparison in spite of the precautions taken in design. For example, one would expect two identical surfaces to appear to have different textures if that under test involved a different metal from that of the comparator. There are also errors arising from imperfections in the master, small changes occurring in the processes of reproduction and variations of texture over the surfaces of the standard and the specimen. Provided care is taken in the visual and tactile estimation, the overall accuracy is to within one-half of a scale step. This is considered a reasonable performance in a simple device for the estimation of surface textures of microwave structures.

(3) ACKNOWLEDGMENTS

The author wishes to thank members of the Engineering Unit, R.R.E., for the considerable amount of work they have done in the measurement of sample finishes and in the preparation of the master scales, and to the Matrix Department, E.M.I. Ltd., Hayes, for assistance in the actual production of the copies.

Acknowledgment is made to the Chief Scientist, the Ministry

Fig. 3.—The final surface-roughness scale.

The copy is in copper with chromium flash on the face. The actual size $4\frac{1}{2}$in × $3\frac{1}{2}$in.

of Supply, and to the Controller of H.M. Stationery Office for permission to publish the paper.

(4) REFERENCES

(1) Harvey, A. F.: "Instruments for use in the Microwave Band," *Proceedings I.E.E.*, Paper No. 1156M, May, 1951 (**98**, Part II, p. 781).

(2) Kuhn, S.: "Calculation of Attenuation in Waveguides," *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 663.

(3) Maxwell, E.: "Conductivity of Metallic Surfaces at Microwave Frequencies," *Journal of Applied Physics*, 1947, **18**, p. 629.

(4) Beck, A. C., and Dawson, R. W.: "Conductivity Measurements at Microwave Frequencies," *Proceedings of the Institute of Radio Engineers*, 1950, **38**, p. 1181

(5) Pippard, A. B.: "The Surface Impedance of Superconductors and Normal Metals at High Frequencies," *Proceedings of the Royal Society*, 1947, **A191**, p. 370.

(6) Morgan, S. P.: "Effect of Surface Roughness on Eddy-Current Losses at Microwave Frequencies," *Journal of Applied Physics*, 1949, **20**, p. 352.

(7) Benson, F. A.: "Waveguide Attenuation and its Correlation with Surface Roughness," *Proceedings I.E.E.*, Paper No. 1467R, March, 1953 (**100** Part III, p. 85).

(8) "Joint Service List of Waveguides—RCL.351" (Ministry of Supply, December, 1950).

(9) "Joint Service List of Waveguide Couplings and Adaptors—RCL.352" (Ministry of Supply, August, 1951).

(10) Reason, R. E., Hopkins, M. R., and Garrod, R. I.: "Report on the Measurement of Surface Finish by Stylus Methods" (Taylor, Taylor and Hobson Ltd., 1944).

(11) "Assessment of Surface Texture," B.S. 1134: 1950 (British Standards Institution).

(12) Beilby, G. T.: "Aggregation and Flow of Solids" (Macmillan, London, 1931).

621.372.8 : 621.357.6

Paper No. 1757R
Mar. 1955

THE ELECTROFORMING OF COMPONENTS AND INSTRUMENTS FOR MILLIMETRE WAVELENGTHS

By A. F. HARVEY, D.Phil., B.Sc.(Eng.), Member.

(*The paper was first received* 14*th April, and in revised form* 21*st August*, 1954.)

SUMMARY

An account is given of the use of electroforming techniques as an accurate and economical method for the manufacture of waveguide components and instruments. Particular reference is made to those for millimetre wavelengths, but some of the techniques can, with advantage, be applied to those for the centimetre wavebands. The development of improved methods and processes, including periodic-reverse-current plating, is described. Methods of manufacture using both permanent and disposable formers are discussed and especial attention is given to the design of the electroformed pieces for ease of subsequent machining. Examples of this technique are given and some electrical and mechanical data are quoted.

(1) INTRODUCTION

Radio-frequency structures for operation within the frequency range 30–300 Gc/s (wavelength 10–1 mm) follow generally those for the centimetric waveband in being made mainly from portions of waveguide transmission lines. Extensive literature[1–6] has been devoted to their design, operation and nomenclature, and a knowledge of this will be assumed. The most important parts are the surfaces carrying radio-frequency currents and the volumes in or around which the magnetic and electric fields exist. In waveguide components and instruments these parts are invariably internal, a fact which gives rise to most of the difficulty of fabrication.

The surface layers, in which the current flows, should be of good conductivity to minimize loss of energy.[7,8] This means that the material must be either pure copper or pure silver and that the surface roughness[9] must be small compared with the skin depth, which for copper at 30 Gc/s (wavelength 10 mm) is only 16 microinches. In practice the surface roughness must be kept[10] less than about 25% or 50% of the skin depth to restrict the increase in attenuation coefficient to 10% and 20% respectively.

The guide wavelength and characteristic impedance of a waveguide depend on its cross-section, and, since waveguide components are generally structures which are either resonant or contain critically spaced admittances, the dimensions of the internal volumes or cavities must be closely controlled. In practice the tolerances required are 1/2 000 to 1/10 000 of the free-space wavelength, and thus range, in millimetre-wavelength components, from 0·0002 in to 0·00001 in.

The majority of such components have hitherto been made by precision machining, and in view of the accuracy demanded such methods will probably continue to play a considerable part in manufacture. There has, however, always been intensive investigation into alternative methods which might supplement or replace those making large demands on limited resources of skilled labour and high-grade machine tools. Such methods include pressure die-casting, hobbing, investment casting, powder metallurgy and electro-deposition. The electro-deposition process has been used successfully both for components and for instruments in the millimetre waveband and will be the only one discussed.

Electroplating techniques have for a long time been applied to waveguide structures as an external plating to improve the appearance and provide resistance to corrosion, as an internal plating to improve surface conductivity and for electrolytic polishing to reduce surface roughness. It is only in the last few years that such techniques have been applied to actual fabrication. The early attempts at electroforming were based on the methods used—under the name "electrotyping"—in the printing and gramophone-record industries. The new methods to be described put the whole technique on a sound engineering basis; in particular, an account will be given of periodic-reversal plating.

In the manufacture of waveguide components by electro-deposition, a former—which is made in the shape of the internal volume required—is electroplated to a sufficient thickness to provide adequate mechanical strength. The former is then removed or withdrawn leaving an electroformed piece or blank. The material for the former may be either a permanent rigid type or it may be expendable. Much attention is given to simplifying the subsequent machining of the electroformed piece, and the electrical characteristics of typical components are described to show that in many instances it has been possible to adopt improved arrangements or designs.

(2) ELECTROFORMING PROCESSES

The requirements which a satisfactory electroforming process has to meet are much more severe than those for normal electroplating. The greater thickness of the deposit, which ranges from 0·040 in to 0·400 in, means that the plating speed should be high and that the throwing power, or ratio of plating speed in areas of low to that in areas of high electric field, must be high and preferably exceed unity. The deposited metal must be homogeneous and free from nodules, cracks and inclusions which may lead to trouble in subsequent machining and expense due to rejection at a late stage in manufacture. In order that the item should possess good mechanical strength and rigidity and yet be easy to machine, the Vickers hardness number* should preferably be in the range 140 to 200.

Many materials can be electroformed, and that chosen for the initial layer on the former is usually silver or copper, to give maximum conductivity. The metal for the subsequent deposit is chosen for mechanical reasons or ease of manufacture and copper is the usual one, although brass or bronze alloys would also serve. Iron would be economical but it is magnetic and very liable to corrosion. Where lightness is required, aluminium or magnesium can be used; where some part is subject to severe wear, a very hard material such as nickel, chromium, or nickel-cobalt alloy can be deposited. When such composite electroforming is used, great care must be taken to obtain thorough bonding of the different metals, and it is desirable that the plated thickness of any one should not be less than

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
Dr. Harvey is at the Radar Research Establishment.

* In the range of hardness considered in the paper the Vickers and Brinell numbers are almost equal.

0·015 in to prevent flaking or peeling during fabrication or use. Consideration must also be given to the different rates of expansion to prevent distortion with temperature changes. As an example of electroforming processes, only copper will be discussed since it is with this metal that most development has been done, the application to other metals being straightforward.

The older electroforming techniques were based on the use of electrolytes containing inorganic salts, usually copper sulphate, the plating being carried out with direct current of density low enough to prevent polarization and burning. The more technical details of these processes are well known and have been extensively described elsewhere;[11] they will not be discussed in the paper.

Components made by these processes were not satisfactory. The metal has a Vickers hardness number of only 60, this softness making machining difficult and the component liable to damage from stresses. The plating speed cannot much exceed 0·001 in/hour, with the result that many components take several weeks to build up. The throwing power is poor so that the article must be removed occasionally from the bath for removal of excessive growth, which process means not only extra work but a risk of damage and difficulty of obtaining a good bond to subsequently plated metal. A further disadvantage is that the plated surface is rough, a property which becomes worse as the deposit thickens so that the formed article shows large nodules, a typical example being given in Fig. 1. In view

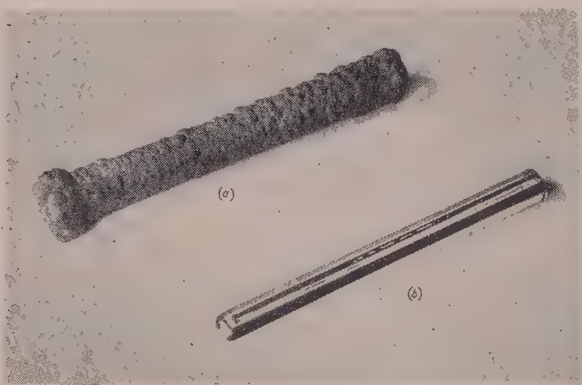


Fig. 1.—Two electroformed pieces in the condition when removed from the plating bath.

The upper piece is by the acid copper process, the lower by the periodic reversal cyanide copper process.

of these disadvantages, intensive development was carried out in order to obtain satisfactory electroforming processes.

One such process is based on an organic electrolyte containing cyanides of copper and other elements used in conjunction with a periodically-reversed plating current. This process[12,13] had been originally developed to give a smooth finished plating; with further work carried out in this country to adapt it for electroforming, it has given very satisfactory results.

With periodic-reverse-current plating, the article being electroformed is made alternately cathodic and anodic, the optimum periods depending on the shape and size but in general being 20–100 and 10–40 sec respectively. The cathodic and anodic currents need not be the same, but best results are obtained when they are nearly so. The anodic cycle must be such as to deplate any unsound metal and nodules while the net current-time product must, of course, be positive. The current densities can be made large so that, although there is a deplating period, the net result is an increased rate of deposition which may be up to 0·006 in per hour without loss of quality. Whilst periodic-reverse-current plating may prove advantageous with the older inorganic electrolytes, a satisfactory process can only be obtained with the cyanide bath mentioned above operating under proper conditions. The best results are obtained with current densities around 80 amp/ft^2 at a temperature of 185°F, although lower current-densities and temperatures down to 120°F can be used with a lower plating speed of about 0·002 in/hour. When the shape of an article is complicated, or when a rough deposit requires smoothing, more sacrificial cycles can be given for short periods.

It is important to keep the solution free from impurities, and a continuously circulating filtration system should be used, the actual filter being, for example, cellulose powder contained in a suitable unit. Because of the high operating temperature and current densities used there is a tendency towards decomposition, with formation of carbonates which cause roughness in the plated deposit. This is overcome by the use, in series with and before the filter, of a purifier containing activated charcoal.

Even and symmetrical deposits are more easily achieved if the solution is agitated, and this is generally done by bubbling air through it. In addition it is desirable to rotate slowly the article being plated so that all parts of it are equally exposed to the anodes, which are, as is usual in plating practice, surrounded with filter bags to retain sludge. The electrolyte is caustic in effect, and to avoid contamination it should come into contact only with such suitable materials as stainless steel, polythene polyvinyl chloride and nylon. The high operating temperature makes fume extraction necessary to remove poisonous cyanide vapours, while there are other small operating details of a chemical and electrochemical nature.

The mechanical properties of the deposited metal are excellent and photomicrographs of sections show that it is free from cracks, treeing and inclusions. The Vickers hardness number depends on the operating conditions and is within the range 120–240. The metal is ductile, free from stresses and machines easily. A typical piece electroformed by this method and in the condition in which it comes from the plating bath is shown in Fig. 1, the smooth exterior, uniform thickness and absence of nodular growths being particularly noticeable.

One of the most fundamental difficulties experienced with electroforming is in obtaining an adequate deposit on concave surfaces. The electric field is always low in such places, at the inside of a sharp corner it vanishes, and, however good the throwing power may be, no deposit will occur there. This shielding effect of the adjacent surfaces becomes worse as the transverse dimensions increase. With the process described above, the thickness of plated metal is fairly constant almost into the corner, the two layers on either side building up with a crack between them. This is shown clearly in Fig. 2, which is a cross-section of the deposit on a cruciform piece of plastic with different radii in all four corners. It will be observed that a continuous layer is deposited until the centre of the circle is reached, from which point a crack develops.

Thus a continuous layer of metal can be achieved only by electroforming in internal corners if some radius, however small, is provided. It will be shown later that this places a limitation on the designs of formers and the methods of electro-deposition. Although a thin continuous layer of metal is sufficient for electrical purposes, considerations of mechanical strength require that this be artificially increased by suitable methods. Those used successfully and shown diagrammatically in Fig. 3 include (*a*) building up with soft solder, amalgams, etc., (*b*) packing[14] with silver powder, (*c*) spraying with metal by the Schori or a similar process, (*d*) placing of suitably shaped inserts, and (*e*) drilling a hole and plugging it. In all these cases electro-deposition

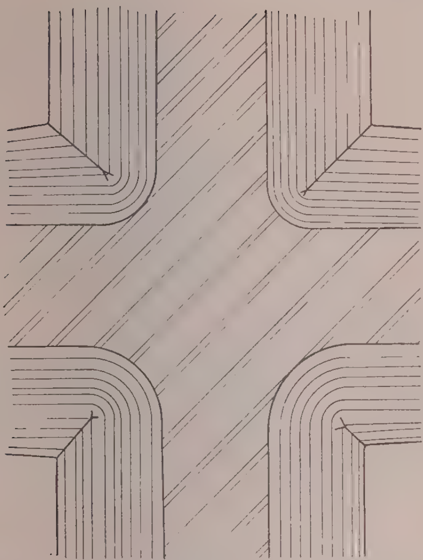


Fig. 2.—Electro-deposition in corners of different radii.

Scale ×10 (approximately). The radii are 0·040, 0·060, 0·100 and 0·120 in respectively.

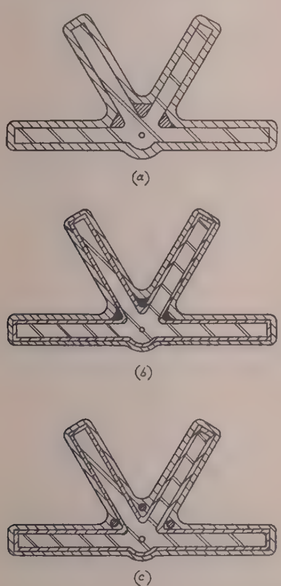


Fig. 3.—Various methods of increasing the thickness of deposit in corners.

(a) Metal inserts shaped to the corner.
(b) Packing the corners by solder, silver powder, amalgams, or sprayed metal.
(c) Drilling and inserting a plug.

is then continued until adequate thickness and strength is obtained.

Electroforming can also be used in conjunction with other techniques such as machining (see Section 5). In another method the metal spraying referred to above can be continued until the component is completed, and if this were carried out in aluminium a saving in weight would be achieved. It is also possible to encase the article, after a thin layer had been deposited, in one of the polyester resins as carried out[15] with potted circuits. There is no reason why several components could not be so encased together to form an equipment sub-assembly which would withstand mechanical shocks and stresses and also adverse climatic conditions.

(3) USE OF PERMANENT FORMERS

In order to use a permanent former it must, of course, be of such a shape that it can be withdrawn after electroforming. The simplest formers are parallel-faced with circular or rectangular sections, and, provided they are properly extracted, as described later, no taper is required. More complicated formers contain two or more sections, the smaller one being at the free end. The transition between differing sections should always be gradual, thereby ensuring that an adequate deposit of metal is obtained.

To prevent distortion and stretching during extraction these permanent formers should be made only from materials such as alloy steels which can be hardened to give high yield point, rigidity and strength. In order to prevent the plated deposit from adhering, the former must be treated by passivation, or coating with a thin layer of a separating medium such as tin.

Formers which are chromium-plated require no subsequent treatment and possess a highly polished long-wearing surface. Other materials which have been used include Invar, molybdenum, titanium, glass, quartz, hard plastics and stainless steel. Stainless steel is very convenient since it can be easily machined and ground to complicated shapes, and it withstands the various chemical processes used in cleaning and electroforming. It cannot be made very hard, however, and the surface is liable in time to become scratched and pitted. Very good surface finishes can be obtained by the use of glass or quartz, but the fragility of these materials restricts their use to short formers of large cross-section.

A series of dimensional measurements have been made, in conjunction with the National Physical Laboratory, on stainless steel formers and pieces electroformed on them. A typical example is a cylinder of length 3 in and diameter $\frac{1}{2}$ in, which was found, on measurement, to be circular and parallel to within 0·00002 in and to have a surface finish of 4 microinches. This former was then plated to a thickness of $\frac{3}{8}$ in, and after extraction the bore of the resulting electroformed piece was carefully measured. The mean diameter was found to be within 0·00002 in of that of the former, the ovality within 0·00005 in, and the surface finish was 5 microinches. Such precise measurements show that an instrument made by this method will have the dimensional accuracy and surface finish of the original former.

A certain amount of machining was then carried out on this electroformed cylinder to provide fittings for the waveguide feeders. On remeasurement the ovality of the bore was found to have increased to 0·00008 in. Earlier versions of this cavity manufactured from metal bar had shown an ovality of as much as 0·0016 in after such external machining, probably owing to the relief of rolling stresses. These results show that an electroformed piece has very little internal stress and is likely to hold its dimensions stable over a long period.

The manufacture of these permanent formers is essentially a problem for the tool-room since the dimensional accuracy required is that usually associated with gauges and similar items. It is preferable to start from bar stock since the rolling process in its production gives a longitudinal structure which tends to maintain the final former straight. After the initial operation of rough turning, milling or grinding, the former is finish-ground to about 0·0002 in oversize. It is then carefully lapped or superfinished to give the final section and surface finish.

Formers of a parallel-faced rectangular section are easily ground in the form of blades by placing direct on the magnetic table of the machine. These blades are then brazed into stainless-

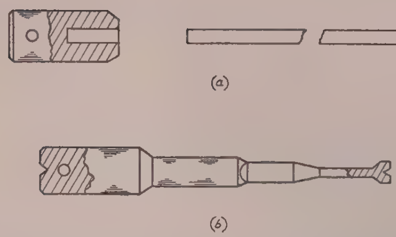


Fig. 4.—Manufacture of permanent formers.

(a) Plain rectangular former made separately and brazed into its head.
(b) Complex former, with various sections, made in one piece. The centre at the free end is removed in a final operation.

steel heads for ease of handling in the electroforming operations, as shown in Fig. 4(a). More complicated shapes are machined from one piece which is held between centres in a fixture which permits accurate setting of the former for each operation. The centre at the free end is removed on completion of the former, as shown in Fig. 4(b).

When chromium-plated formers are required, a core is first made from a steel which can be hardened to about 600 Vickers without any distortion. An allowance of 0·006in per face is made for plating and any edges are given a $\frac{1}{64}$in radius. The core is then hard chromium-plated to a thickness sufficient to allow final grinding and lapping as before.

In order to provide a square end-face on the electroformed piece, a stop made of polythene is usually put on the mandrel to form a watertight seal. The head of the former is insulated by a polythene tube of sufficient length to come well above the surface of the plating bath. Electro-deposition is usually carried out over the whole surface of the former so that it is protected from dust or other foreign particles. If the closed end is not required, it can easily be removed subsequently. When the cross-section of the former is very small an extra portion, usually cylindrical, is incorporated to serve as a guide during extraction and to provide an adequate support during machining of the electroformed piece. This portion is finally discarded; Fig. 5 shows an example of such electroforming.

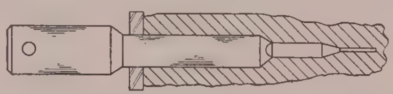


Fig. 5.—Permanent former with electroformed piece in position.

The polythene stop which gives a square face is shown. This former has an extra cylindrical portion to serve as a guide during extraction and as a support for the electroformed piece in subsequent machining.

As a preliminary to plating, the former is thoroughly cleaned by chemical and electrolytic methods. It is then covered with a thin layer of silver (or copper, if required) by plating in a very weak electrolyte termed a "strike bath." The former is finally plated in the baths previously described to give the required thickness. For most millimetre-wavelength components a 0·015in deposit of silver is given, followed by copper. Thinner deposits of silver are found to lead to peeling during subsequent machining; thicker deposits lead to undue cost in the larger sizes.

On completion of plating, the former, with the electroformed piece surrounding it, is removed from the bath, washed and prepared for extraction. With short cylindrical sections, simple immersion in hot water enables the former to be pulled out by hand. With the more complicated and longer formers the extraction has to be done with considerable care and it is advisable to use a special machine, such as the one illustrated in Fig. 6, in which the extraction is done vertically while the electroformed piece is immersed in water or oil at about 200°F. The pull

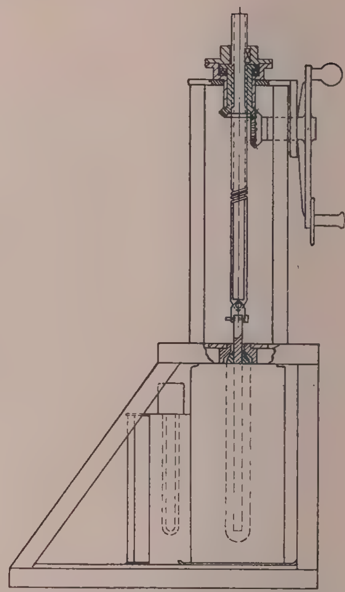


Fig. 6.—Machine suitable for extracting formers.

The electroformed piece, whilst immersed in the hot liquid, is held against a self-aligning stop. Slow rotation of the handle extracts the former with a straight pull.

required is such that the elastic limit of the material of the former sets an upper limit to the length which can be electroformed. Thus a 0·280in $\times$ 0·140in section would have a maximum length of about 10in, whereas one of only 0·034in $\times$ 0·017in section is limited to 1in.

Waveguide components involving coupling apertures and slots can be electroformed with formers provided with extra pieces, which should be of plastic to prevent shielding in the corners. These pieces are cheap and are discarded after use. Fig. 7 illustrates two such formers.

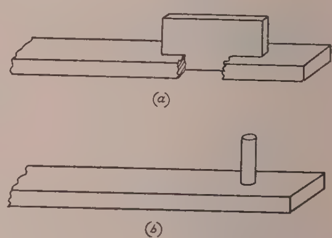


Fig. 7.—Formers for the electro-deposition of waveguide sections with coupling apertures.

The strip in (a) and peg in (b) are made of plastic to prevent shielding in the corners thus formed.

Where there is more than one waveguide section, as for example, in T-junctions, it is usual to employ multiple mandrels. The alignment is important and is ensured by supporting them in jigs, which can also contain means of supporting any inserts required. These inserts are made of metal, ultimately forming an integral part of the finished article, and are used to strengthen corners or provide coupling apertures between the waveguide sections.

Some care is necessary in the choice of materials for such jigs. Plastic materials tend to warp in or contaminate the plating solutions and steel tends to corrode in time. Aluminium is light but again not satisfactory in caustic electrolytes, and thus it is better to use stainless steel, which is entirely suitable despite

its high cost. To prevent deposition on the jigs they are covered all over with a stopping-off wax or tape.

Examples are given in Fig. 8 of jigs for three different types of waveguide component. For ease of electroforming the hybrid ring arrangement includes two extra mandrels, the arms so formed being later plugged. One disadvantage of these multiple-mandrel assemblies is that the junctions are not perfect and with the improved plating process a deposit is thrown inside even very closely fitting joints. This causes a small amount of flash which has to be removed during subsequent machining. Attempts to bridge these gaps by depositing, at an early stage, from coarse-grain plating baths such as nickel, have been only partially successful.

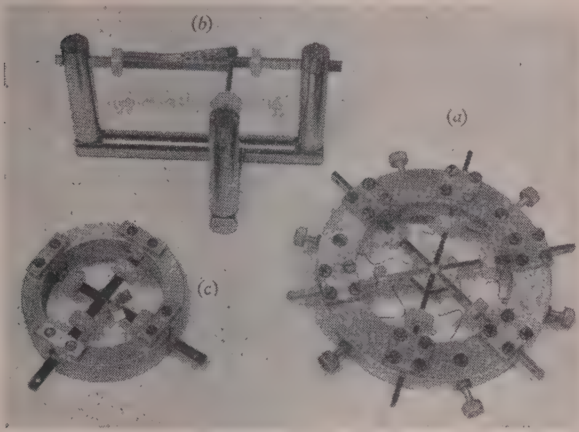


Fig. 8.—Various types of jig for supporting steel formers.

(*a*) Multiple-hole directional coupler. (*b*) Cross-over directional coupler. (*c*) Hybrid ring, two of the arms being finally plugged.

(4) USE OF DISPOSABLE FORMERS

With disposable formers, withdrawal imposes no restrictions and thus this type is very suitable for electroforming complex shapes. Although such formers can be accurately fabricated, they are thus used only for the electroforming of a single component for experimental purposes. The application which will be mainly considered here is that in which the formers are cast or moulded in large quantities.

One of the simplest materials to use for these formers is wax, a suitable type being Seekay because of its high melting point of about 100°C, low shrinkage on setting and narrow plastic range, being very fluid when just above the melting point. These waxes are sometimes loaded with materials which increase their strength and reduce shrinkage on cooling. More accurate formers can be obtained by the use of air pressure and possibly centrifugal casting. The wax is removed by melting, and since it can be used again the process is economical.

Another suitable class of materials are alloys such as Cerromatrix, which has the useful property that it does not shrink on solidifying and Mazak, which can be accurately cast. Such materials can be removed partly by melting, but, since there is a tendency towards amalgamation with the electroformed metal, they usually require caustic solutions to remove all traces. In some cases aluminium can be used as the former, and here again removal is by means of caustic soda. These metals are also convenient when formers are individually machined for electroforming only a few components.

Plastic materials have been used with success since they can be injection-moulded under high pressure to give, in properly designed moulds, formers of acceptable accuracy. Diakon, one of the methacrylate resins, has been found suitable because of its small shrinkage on moulding, low water absorption and ease of removal in, say, a suitably designed Soxhlet apparatus using a solvent such as chloroform. Such plastic material is economical for millimetre-wavelength components but not for much larger sizes, since it cannot be used again.

The moulds for these plastics must be strongly made since injection pressures of about 30 000 lb/in^2 are customary. An allowance for shrinkage which, in Diakon, for example, is about 1%, can be made in the relevant dimensions of the mould. To ensure that the shrinkages of all formers are identical, the moulding is preferably carried out on an automatic machine in which the temperature of the mould, the injection and cooling times and other conditions are maintained constant. By such means the dimensions of a waveguide section of, for example, $0\cdot280$ in $\times$ $0\cdot140$ in can be kept within $\pm 0\cdot0005$ in. Some examples of plastic mouldings for a hybrid ring, a hybrid-T and a complex duplexer unit are illustrated in Fig. 9.

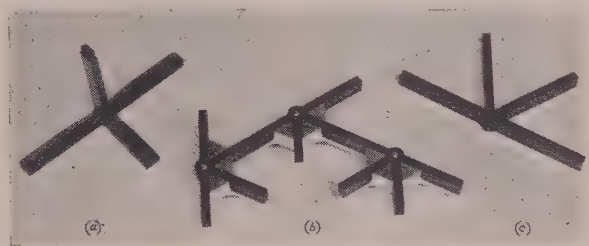


Fig. 9.—Disposable formers made in plastic.

(*a*) Hybrid-tee. (*b*) Multiple-ring duplexer unit. (*c*) Hybrid ring.
The section of the waveguide is $0\cdot280$ in $\times$ $0\cdot140$ in; the presence of a radius in the various corners should be noted.

The mould is usually built up from separate pieces, whose faces are lapped so as to give surface finishes of about two microinches. In order to ease the subsequent electroforming, the mould must be constructed so that all internal corners on the former are provided with the maximum permissible radii. To facilitate the removal of the former, the mould is usually divided into two halves and it is important that these should mate closely to avoid flash. Some care must be given to the point of injection since the viscous plastic does not flow so readily into the more remote regions; otherwise undersize formers will result. With hybrid junctions a symmetrical injection is preferred, although this means that the surplus plastic left in the injection gate must be carefully removed. It is generally easy to arrange for ejector pins by slightly increasing the length of the arms.

These waxes and plastics are insulators, and the surfaces of the former have to be made conducting before plating can take place. A conducting layer can be provided by spraying with Aquadag which is subsequently removed, but the poor surface finish and tolerances restrict this method to the larger sizes. With the plastic formers, the usual method is to immerse them in a chemical silvering solution of conventional composition. This silvering can more easily be carried out, especially when large quantities or sizes are required, by a gun devised by the Printing and Allied Trades Research Association (P.A.T.R.A.). This contains a nozzle from which two jets of the constituent solutions emerge and, on coalescence, form a thin adherent silver film on the former. Metal inserts can be included in the assembly provided they are of silver or are silver plated. Another method of giving a conducting layer is by evaporation under low pres-

sure, which permits a variety of metals to be used, although the coating of complex or re-entrant shapes will necessitate rotation of the former or the use of multiple cathodes.

Most of these disposable formers are liable to distortion in hot plating solutions, Diakon, for example, starting to soften around 120°F, and above 140°F serious swelling and distortion is certain. The effects of this can be minimized by leaving the ends of the electroformed piece open and also by holding, in a locating jig, those ends whose position is important. If a reduced speed of plating of around 0·002 in/hour can be tolerated, it is better to keep the electrolyte at 120°F.

Any inserts employed can be supported direct by the former or held in position by a locating plate or jig. The shape of these inserts should be such that sharp corners, which cause cracks in the deposit, are avoided. Thus a post or pin should be shaped as in Fig. 10(*c*) to secure sound bonding and deposition.

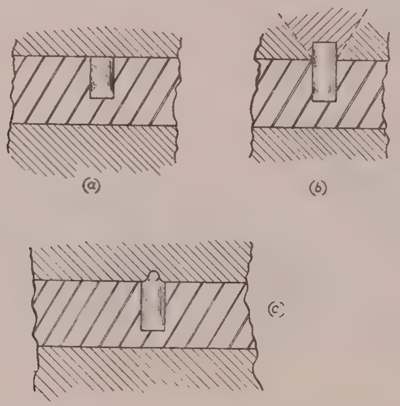


Fig. 10.—Methods of electroforming inserts in position, using disposable formers.

(*a*) The flush-fitting pin gives sound deposition but poor keying.
(*b*) The protruding pin gives poor deposition but good keying.
(*c*) The shaped protruding pin gives sound deposition and good keying.

These plastic formers can also be joined together to form complex assemblies, the whole unit then being electroformed. This is more economical than the production and assembly of the individual items. An example of such a complex assembly has been given in Fig. 9. In this particular instance, the plastic formers are supported on a properly shaped base-plate. This prevents distortion, supports the inserts and ultimately forms one face of the electroformed piece.

(5) FABRICATION OF ELECTROFORMED PIECES

Waveguide components and especially instruments are rather complex and electroforming of the complete item is rarely possible. Thus some final machining and assembly is usually required, although in the simpler components this can be restricted to the assembly of connecting flanges and arrangement of mounting holes.

The basic electroformed piece is a short straight section of rectangular waveguide which, by being produced on an accurate former, can have a tolerance of 0·0002 in as compared with the 0·001 in of drawn tubing. After removal of the former this precision waveguide, as it is termed, is supported between centres and machined externally to have a cylindrical[16] outer section which is concentric with the rectangular bore. This is done so that the fitting of connecting flanges, which now have a cylindrical instead of the conventional rectangular hole, is made easy and accurate. Angular alignment of the flange is ensured by a simple soldering jig. This guide is made with various external diameters so that it can be inserted, as a separate piece, into instruments in which locating bores have been machined. Precision waveguides of cylindrical internal section can also be readily electroformed and machined.

It is often desirable to have a component coupling waveguides of two different sections; this can easily be achieved by electroforming on a suitably tapered former. In most cases a permanent former can be used, but when transforming from cylindrical to rectangular waveguide it is sometimes necessary to use a disposable type.

Many waveguide instruments incorporate micrometer heads and the former can be shaped so that the housing for the micrometer bearing is electroformed in position. Thus an accurately made former can produce pieces which, without any elaborate machining, ensure proper fit of the micrometer and its alignment[17] with the waveguide section.

These electroformed pieces require machining in order to make them the correct external shape and size, to fit connecting flanges, to accept or mate with other parts and sometimes to improve the appearance. It is undesirable to use the original electroforming former as a support during such machining since it is not made for this purpose and is too expensive to be subjected to risk of damage. It is better to use a special mandrel which can be of the proper shape, with suitable clearance, made of hardened steel to reduce distortion and provided with means for centring and driving. In sections of, say, 0·100 in × 0·050 in and below, these mandrels become fragile, and it is desirable to include in the electroformed piece an extra portion, adequate to withstand machining stresses, which is finally removed. The walls of these millimetre-wavelength components can be made thick to provide strength and rigidity without excessive weight. In fact, the requirements of accuracy often mean that items such as micrometer heads, dial gauges, etc., are much larger than the purely waveguide portions.

Local parts which are liable to wear—such as V-slides and other moving pieces—or which require exceptional strength, can be electroformed from such hard materials as nickel, chromium and nickel-cobalt alloy. The throwing power and speed of deposition of these high-valency materials is low and their use should be restricted to essential places. Items which must be very light in weight can be electroformed in magnesium or aluminium,[18–20] but since such processes are at present only on a small or laboratory scale, their commercial use is not so advanced as that of other metals.

The assembly of the pieces forming the electroformed article is best done by soft soldering, which avoids the distortion inherent in high-temperature brazing. Sometimes it is possible to eliminate a soldered joint by further electroforming. Thus a hybrid ring might be made by machining channels in a block of copper, temporarily filling these with wax or plastic and then electroforming the fourth side.

(6) ELECTRICAL DESIGN AND CHARACTERISTICS

The final electroformed article will fall into one of two classes. The first includes instruments used for measuring purposes; in view of the accuracy required they are usually made on permanent formers. The second includes components which only serve as subsidiary parts of equipment; these can usually be satisfactorily electroformed on mass-produced disposable formers. The development of the electroforming techniques has enabled existing components and instruments to be redesigned with improved performance and for ease of manufacture; it has also made possible new types which previously could be made only with difficulty.

There are many uses to which the precision waveguide described

in the previous Section can be put. In the form of inserts in instruments and components it can be likened to elements such as lenses and prisms used in optical instruments which form units separate from the purely mechanical parts. Such guides have been made with internal sections from 0·280in × 0·140in to 0·034in × 0·017in, the smaller sizes being available only in quite short lengths of one inch or so. A range of these guides have recently been standardized[21] (see technical data in Table 1).

When so made, not only do these units possess an accurate bore but also there is proper alignment of the micrometer and waveguide axes. Fig. 14 illustrates a plunger unit for waveguide size 0·0510in × 0·0255in, the movement being by means of a micrometer head with 13mm travel. Furthermore, wavemeters with such rectangular cavities are free from degenerate and higher modes and are much preferred for general purposes to the usual H_{11} (TE_{11}) cylindrical-mode type.

Table 1

STANDARD SIZES OF PRECISION ELECTROFORMED WAVEGUIDE

R.C.S.C. type number	Inside rectangle		Cut-off for fundamental mode		Recommended operating range		
	Width	Height	Frequency	Wavelength	Frequency	Wavelength	Wave number
	in	in	Gc/s	mm	Gc/s	mm	cm^{-1}
WG22P	0·2800	0·1400	21·1	14·23	26·5–40·0	11·3–7·5	0·88–1·33
WG23P	0·2240	0·1120	26·3	11·38	33·0–50·0	9·1–6·0	1·10–1·66
WG24P	0·1880	0·0940	31·4	9·56	40·0–60·0	7·5–5·0	1·33–2·00
WG25P	0·1480	0·0740	39·9	7·52	50·0–75·0	6·0–4·0	1·66–2·50
WG26P	0·1220	0·0610	48·4	6·20	60·0–90·0	5·0–3·3	2·00–3·00
WG27P	0·1000	0·0500	59·0	5·08	75·0–110·0	4·0–2·7	2·50–3·70
WG28P	0·0800	0·0400	73·8	4·06	90·0–140·0	3·3–2·2	3·00–4·55
WG29P	0·0650	0·0325	90·9	3·30	110·0–170·0	2·7–1·8	3·70–5·55
WG30P	0·0510	0·0255	115·8	2·59	140·0–220·0	2·2–1·4	4·55–7·13
WG31P	0·0430	0·0215	137·5	2·18	170·0–260·0	1·8–1·2	5·55–8·33
WG32P	0·0340	0·0170	173·5	1·73	220·0–325·0	1·4–0·9	7·13–11·10

Precision waveguides electroformed to the larger diameters can be machined to form bearings for sliding parts, and locations for side arms in T-junctions, cross-over connections and directional couplers, without the complication of strengthening supports. An example of a multiple-hole-type directional coupler using 0·148in × 0·074in precision waveguide is given in Fig. 11, and of a matching unit using 0·280in × 0·140in guide in Fig. 12.

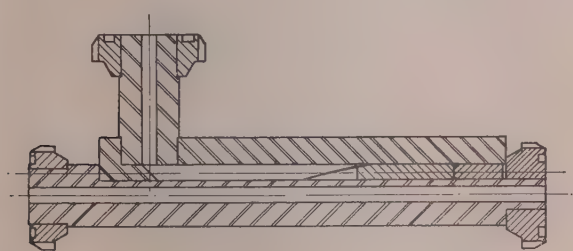


Fig. 11.—Multiple-hole directional coupler.

The example illustrated uses a waveguide of internal section 0·148in × 0·074in and is intended for use in the 4–6mm waveband. The coupling is by means of 12 holes, but other apertures can be used.

Electroforming is very suitable for the manufacture of taper transitions from one waveguide size to another; Fig. 13 shows such a taper from a section of 0·122in × 0·061in to one of 0·034in × 0·017in. Horn flares and three-sided corner reflectors, provided they are not too large, can also be electroformed. Bends and corners, whose internal section varies along their length in order to give optimum impedance matching, can also be readily made by such techniques. This also applies to sections tapering in two dimensions or when combined with bends and twists. In these cases disposable formers must, of course, be used.

Electroforming has made relatively easy the manufacture of plunger units, wavemeters and stub tuners in which a cylindrical plunger, actuated by a micrometer screw-head, moves in a rectangular waveguide propagating in the H_{01} (TE_{01}) mode.

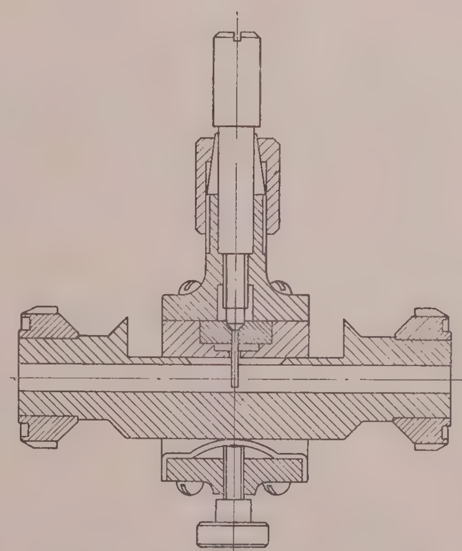


Fig. 12.—Moving probe matching assembly.

The main body has an internal section of 0·280in × 0·140in, and the electroformed thickness is sufficient to allow machining of the locating surfaces of the sliding carriage.

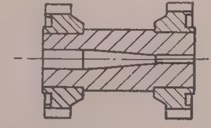


Fig. 13.—Taper adaptor.

The example illustrated incorporates a section ⅜in long, tapering from 0·122in × 0·061in to 0·034in × 0·017in.

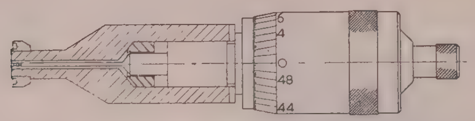


Fig. 14.—Micrometer plunger unit.

The example illustrated has a 0·080 in × 0·040 in section and is suitable for use in the 2–3 mm waveband. If the free end is closed and fitted with suitable input and output connections, a general purpose wavemeter is obtained.

Electroforming on permanent cylindrical formers is the only method at present capable of the quantity production of the very accurate bores required for standard wavemeters and cut-off attenuators in the millimetre wavebands. The conductivity of the electroformed metal is close to that of the pure solid, and this, together with the possibility of obtaining surfaces of low roughness value, enables resonant cavities of high Q-factor to be obtained. A typical wavemeter, such as that illustrated in Fig. 15,

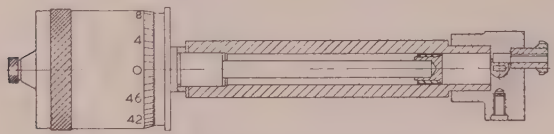


Fig. 15.—H_{01} cylindrical mode wavemeter.

The cavity is cylindrical and parallel to within 0·00002 in. The diameter is 0·6831 in so that its cut-off frequency is that of 0·280 in width rectangular waveguide. The instrument is suitable for use from 8 to 10 mm wavelength.

with a cavity of 0·6831 in diameter, resonating in the H_{01} (TE_{01}) cylindrical mode at 9 mm wavelength, has a Q-factor ranging from 40 000 at the 12th half-wave resonance to 50 000 at the 24th. The accuracy of wavelength measurement is such that, during some initial work on these techniques and using a known frequency standard, Mr. C. R. Ditchfield in unpublished work obtained a value for the velocity of light of $2 \cdot 99790 \times 10^{10}$ cm/sec, which is very close to the usually accepted value. Such electroformed cylindrical cavities have also been used at this wavelength for the determination of dielectric constant and loss tangent of low-loss materials, as reference cavities in frequency-stabilizing systems and as noise filters.

The electroforming in position of inserts enables directional couplers to be made with a variety of coupling apertures. In the example given in Fig. 8, the insert takes the form of a thin metal strip with correctly spaced and sized holes. Inserts with apertures of improved directional properties, such as T-shape or cruciform, can be made of thin metal by photo-etching methods and electroformed in position between either permanent or disposable formers. The multiple assembly already seen in Fig. 9 is for electroforming a transmitter-receiver duplexer circuit using hybrid rings for the local oscillator feeds and for the signal and a.f.c. balanced mixers. Measurements on such a unit gave a loss of about 0·5 dB and a discrimination between the arms not worse than 40 dB over a 3% bandwidth. Other arrangements incorporating hybrid rings or tees and directional couplers could be similarly made.

Electroforming can also be applied to the manufacture of flexible waveguides by using a disposable former of the required cross-section but with a series of transverse grooves.[22] These may be a half-wavelength deep or alternatively very shallow and closely spaced, so that a corrugated waveguide, operating in the pass band,[23] is produced. This former is then electroformed to about 0·005 in thickness along its length, but rather more at the end for fitting of the flanges. A rubber jacket is finally moulded around the guide to give it mechanical support. Such flexible guides appear more promising in the very small sizes than those of the usual soldered convolute construction.

It is often necessary, when weight and space saving is important, to fill the waveguide with a dielectric material, thus enabling the dimensions to be reduced by the square root of the permittivity. Since the former remains in position after electro-deposition the material must be such that it has a small loss tangent and can provide a good bond with the deposited metal which does not fail under temperature cycles.

The fact that electroforming techniques have enabled waveguides of very small section to be made has enabled a small range of components and instruments for operation at wavelengths of 1 mm to be constructed. At such short wavelengths semi-optical techniques[24] can be used, since the associated lenses, reflectors and prisms become of reasonable size. Thus the technique of waveguide manufacture enables them to be used over their whole usable wavelength range.

(8) ACKNOWLEDGMENTS

The author is indebted, in the development of these techniques, to much assistance received from the Engineering Unit, R.R.E., Messrs. Cruickshank Ltd., Hilger and Watts Ltd., Evered and Co. Ltd., and Projectile Engineering Co. Ltd. Acknowledgment is made to the Chief Scientist, Ministry of Supply, and to the Controller, H.M. Stationery Office, for permission to publish the paper.

(9) REFERENCES

(1) Harvey, A. F.: "Instruments for use in the Microwave Band," *Proceedings I.E.E.*, Paper No. 1156M, May, 1951 (**98,** Part II, p. 781).
(2) Stratton, J. A.: "Electromagnetic Theory" (McGraw-Hill, New York, 1941).
(3) Ragan, G. L.: "Microwave Transmission Circuits" (McGraw-Hill, New York, 1948).
(4) Montgomery, C. G.: "Technique of Microwave Measurements" (McGraw-Hill, New York, 1947).
(5) Montgomery, C. G.: "Microwave Circuits" (McGraw-Hill, New York, 1948).
(6) Ramo, S., and Whinnery, R.: "Fields and Waves in Modern Radio" (Wiley and Sons, London, 1944).
(7) Maxwell, E.: "Conductivity of Metallic Surfaces at Microwave Frequencies," *Journal of Applied Physics*, 1947, **18,** p. 629.
(8) Beck, A. C., and Dawson, R. W.: "Conductivity Measurements at Microwave Frequencies," *Proceedings of the Institute of Radio Engineers*, 1950, **38,** p. 1181.
(9) Harvey, A. F.: "A Surface Texture Comparator for Microwave Structures" (see page 219).
(10) Morgan, S. P.: "Effect of Surface Roughness on Eddy Current Losses at Microwave Frequencies," *Journal of Applied Physics*, 1949, **20,** p. 352.
(11) Blum, N., and Hogaboom, G. B.: "Principles of Electroplating and Electroforming" (McGraw-Hill, New York, 1949).
(12) Jernstedt, G. W.: "Periodic Reversal Copper Plating," *Proceedings of the American Electroplaters Society*, 1949, **36,** p. 63.
(13) Westinghouse Co.: British Patent No. 630165.
(14) Balmer, J. R.: British Patent No. 685247..
(15) Dummer, G. W. A., and Johnson, D. L.: "Printed and Potted Electronic Circuits," *Proceedings I.E.E.*, Paper No. 1407R, November, 1952 (**100,** Part III, p. 177).
(16) Harvey, A. F.: British Patent No. 691128.
(17) Harvey, A. F.: British Patent No. 683246.
(18) Couch, D. E., and Brenner, A.: "A Hydride Bath for the Electrodeposition of Aluminium," *Journal of the Electrochemical Society*, 1952, **99,** p. 234.
(19) Safranek, W. H., Schickner, W. C., and Faust, C. L.: "Electroforming of Aluminium Waveguides using Organo-aluminium Plating Baths," *ibid.*, p. 53.
(20) Hurley, F. H., and Weir, T. P.: "Electrodeposition of Aluminium from Non-aqueous Solutions at Room Temperature," *ibid.*, 1951, **98,** p. 207.
(21) "Joint Services List of Waveguides—RCL351" (Ministry of Supply, December, 1950).
(22) Lines, A. W., and Balmer, J. R.: British Patent Application No. 30360/52.
(23) Lines, A. W., Nicoll, G. R., and Woodward, Mrs. A. M.: "Some Properties of Waveguides with Periodic Structure," *Proceedings I.E.E.*, Paper No. 941R, July, 1950, (**97,** Part III, p. 263).
(24) Culshaw, W.: "The Michelson Interferometer at Millimetre Wavelengths," *Proceedings of the Physical Society*, 1950, **63,** Part B, p. 393.

621.396.11.029.6

Paper No. 1780 R
Mar. 1955

AN EXPERIMENTAL STUDY OF THE PROPAGATION OF 10 CM RADIO WAVES OVER A SHORT NON-OPTICAL SEA PATH

By E. F. STACK-FORSYTH, B.Sc., Ph.D., Associate Member.

(*The paper was first received* 24*th May, and in revised form* 11*th October*, 1954.)

SUMMARY

A series of tests in microwave propagation over a sea path of length 1·14 times the optical horizon was conducted off the coast of Natal during the winter months, April–August, 1953. The purpose of the investigation was to study the effect upon signal strength of variations in the structure of the refractive-index profile of the atmosphere in the first few hundred feet above sea level. The radio equipment operated at a wavelength of 10cm, and vertical polarization was used throughout the tests.

LIST OF SYMBOLS

a = Radius of earth, ft.
d = Duct height, ft.
x_∞ = Lapse rate of modified refractive index above duct.
m = Profile index.
M = Refractive modulus, M-units.
n = Refractive index of atmosphere.
N = Modified refractive index of atmosphere.
P = Transmitter power, watts,
p = Gradient of modified refractive index below duct minimum.
q = Gradient of modified refractive index above duct minimum.
z = Height above surface of earth, ft.

(1) INTRODUCTION

It is well known that the problem of propagation of short electromagnetic radiation round a spherical earth can be reduced to that of propagation over a plane earth by adoption of the artifice of using a modified refractive index of the atmosphere, which is related to the actual refractive index by the formula

$$N = n(I + z/a) \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Under normal conditions the gradient of N is positive and numerically equal to about $0\cdot036 \times 10^{-6}$ per foot. Under abnormal conditions, when there is a high lapse rate of water-vapour pressure or a temperature inversion in the atmosphere, this gradient may be less or may even be negative over a limited height interval of a few hundred feet, becoming normal at greater altitudes. Under these conditions, electromagnetic radiation will be refracted towards the earth, with the result that ranges considerably in excess of the optical horizon may be obtained. The radiation then behaves as if it were trapped in a duct close to the surface of the earth. It can be shown that the degree of trapping depends both upon the frequency of the radiation and the height over which the gradient of modified index is negative, and is such that only microwaves will be appreciably affected by ducts of the heights generally encountered.

The mathematical theory of duct propagation has been developed by Booker and Walkinshaw,[2] and by Freehafer and Furry.[1] Booker and Walkinshaw considered a modified index profile represented by

$$N^2 = 1 - 2x_\infty[z - (d^{1-m}z^m/m)] \quad . \quad . \quad . \quad (2)$$

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
Dr. Stack-Forsyth is at the Natal Technical College, Durban, Natal.

Taking values of $\frac{1}{2}$ and $\frac{1}{5}$ for m, the profile equation becomes

$$\textit{Square-law.} \quad M = 4 \times 10^{-2}[z - 2\sqrt{(zd)}] \quad . \quad . \quad . \quad . \quad (3)$$

$$\textit{Fifth-root law.} \quad M = 4 \times 10^{-2}[z - 5(zd^4)^{\frac{1}{5}}] \quad . \quad . \quad . \quad . \quad (4)$$

where

$$M = (N - 1) \times 10^6 \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Furry dealt with a profile consisting of two linear portions of gradients p and q, below and above the duct minimum respectively, and calculated field strengths for the case where

$$s^3 = p/q = -1, \text{ and } q = 0\cdot036 \times 10^{-6} \text{ per foot.}$$

Fig. 1 shows the various profiles for duct heights of 50, 100, 150 and 200ft. It was the purpose of the investigation to attempt to correlate the signal strengths obtained in practice with those predicted on the basis of these three profiles.

(2) EXPERIMENTAL PROCEDURE

(2.1) Propagation Path

Fig. 2 is a map of the locale of the tests. The path was entirely over the sea, 21·1 statute miles in length, and between terminals at altitudes of 47ft above mean sea level. It was approximately parallel to a homogeneous coast line. The prevailing winds blow approximately parallel to the coast line, i.e. south-west and north-east.

(2.2) Transmitter

The transmitter was a type CV35 klystron giving a rated power output of 100mW. It was modulated at 1000c/s by means of a square wave applied to the reflector electrode. The associated power supplies were fully regulated, and radiation was from a dipole placed at the focus of an 18in-diameter paraboloidal mirror. Constancy of power output was checked by noting the output from a radiation monitor placed near the antenna; this did not vary by more than 2·5dB from the mean value throughout the series of tests.

(2.3) Receiver

The receiver consisted of a crystal mixer followed by a 90dB six-stage i.f. amplifier. Stagger tuning was employed to obtain a bandwidth of 5Mc/s, centred on 21·7Mc/s. A small amount of automatic gain control was applied to the first three stages in order to obtain the desired overall characteristic. A large bandwidth was chosen so as to avoid errors in field-strength measurement owing to variation in frequency of the transmitting klystron and of the local oscillator. A two-step attenuator was incorporated in the receiver so as to give two ranges of signal input, i.e. 0–100μV and 0–1 000μV input to the i.f. amplifier. The receiver operated a recording milliammeter. A d.c. crystal checker was built into the receiver so that a rough check on the crystal characteristic could be made frequently. This checker enabled the forward resistance to be measured at an applied voltage of 0·25 volt, and the back resistance at an applied voltage of 1 volt. There was no significant change in the values of these

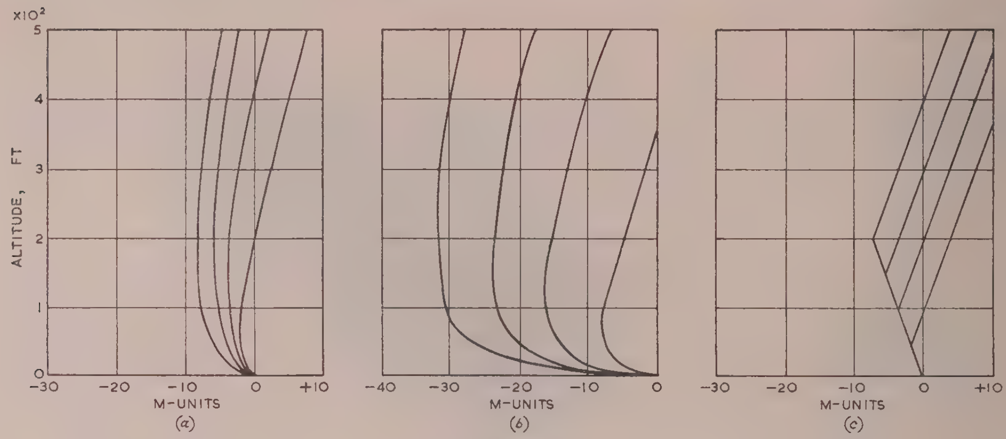


Fig. 1.—Theoretical *M*-profiles.

(*a*) Square law. (*b*) Fifth root. (*c*) Bilinear.

throughout the series of tests. Provision was made in the receiver for altering the coupling between the local oscillator and mixer, so that the crystal could readily be set to operate always on the same point of its characteristic, by noting the d.c. crystal current. The amplification characteristic of the receiver between input to the i.f. amplifier and output to the recorder was checked before and after the series of tests; the results did not differ by more than 3 dB over the range of signal strengths encountered.

(2.4) Atmospheric-Sounding Equipment

Temperature and humidity of the atmosphere were measured by means of wet and dry thermistors. These were mounted in a double-walled radiation shield provided with a conical top to eliminate radiation from above, and were airborne by means of balloons or kites. Their resistance was measured by means of a Wheatstone bridge operated off balance and connected to the airborne unit by a special light-weight cable. In calm weather, altitudes of approximately 600 ft were attained; when windy the kite reached an altitude of about 300 ft. The reading accuracy of the equipment was 0·2°C, but it was estimated that in actual practice the accuracy was not better than 0·5°C.

It was only possible to make soundings at the northern (receiving) end of the link, but since the propagation path was roughly parallel to a homogeneous coastline, and since the prevailing winds were roughly parallel to this path, it was felt that the results obtained in this way would be a fair approximation to the average conditions over the whole path. There were occasions when a time lag of up to about one hour existed between a change in signal strength and a change in atmospheric conditions, owing to a wind springing up or dying away. When the wind was from the north-east, the change was noted first in atmospheric conditions, and when from the south-west, a change in signal occurred first. There was always an element of doubt as to whether the soundings made at the coast gave the same results as would be obtained over the sea. In taking a sounding, observations were made both on ascent and on descent of the unit, the average time taken for a sounding being half an hour.

(2.5) Calibration of the Link

An estimate of the free-space signal level at the receiver was made as follows: Megaw[3] concluded, as a result of observations made of signal strengths at short ranges over the sea, that, at points well within the optical horizon, the field strength was always within ±2 dB of the value calculated on the basis of a standard atmosphere. Accordingly the receiver was set up at two points, shown in Fig. 2, well within the optical horizon, and the signal strength determined in terms of the output from the receiver. The actual field strength to be expected at these points due to a dipole radiating 1 kW was then calculated by a method described by Norton,[4] using an effective earth radius of 5 280 miles, i.e. for standard atmospheric conditions. For the Beachwood site (range, 6·8 miles), the measured receiver output was 7·2 mA, which corresponded to an i.f. amplifier input of 32 dB above 1 μV.

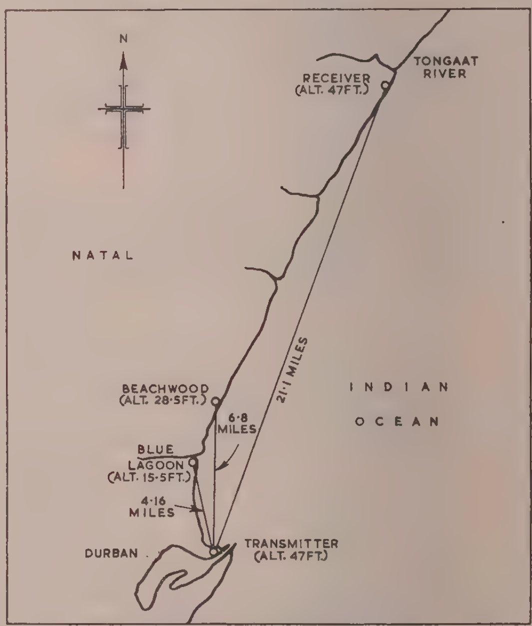


Fig. 2.—Map of propagation path.

The calculated signal strength at a range of 6·8 miles for a dipole radiating 1 kW was found to be 18 mV/m. If the actual transmitter had been equivalent to a dipole radiating *P* watts, the theoretical signal strength at Beachwood would have been $18\sqrt{(P/1\,000)}$ mV/m.

Under free-space conditions the signal would have this value at a range of 7·67 miles. Thus the free-space signal input to

the i.f. amplifier at 7·67 miles was 32 dB above 1 μV. At Tongaat (range, 21·1 miles) the free-space signal input to the i.f. amplifier would be 8·8 dB below this, i.e. 23·2 dB above 1 μV. From the receiver calibration curve it was found that a signal of this value would produce an output of 1·4 mA.

A similar calculation for the Blue Lagoon site gave a value of 1·0 mA. These two results differ from their mean, i.e. 1·2 mA, by less than 2 dB, and so this value was taken as representing the free-space signal level at the receiver.

An element of uncertainty in the result arises from the fact that the receiver was only calibrated from the input to the i.f. amplifier. However, it is reasonable to assume that the mixer was linear over the range of signal strengths encountered.

(2.6) Operational Procedure

It was not possible to conduct the tests continuously, but as fas as possible the operating times were arranged to cover those periods of the day when changes in conditions were most prevalent, i.e. during the early morning and late afternoon. A total of 37 tests was made, covering a total operational period of 170 hours. Altogether 85 atmospheric soundings were made.

(3) EXPERIMENTAL RESULTS

Typical signal records and atmospheric soundings are presented in Figs. 3–5. These show the variation in signal strength,

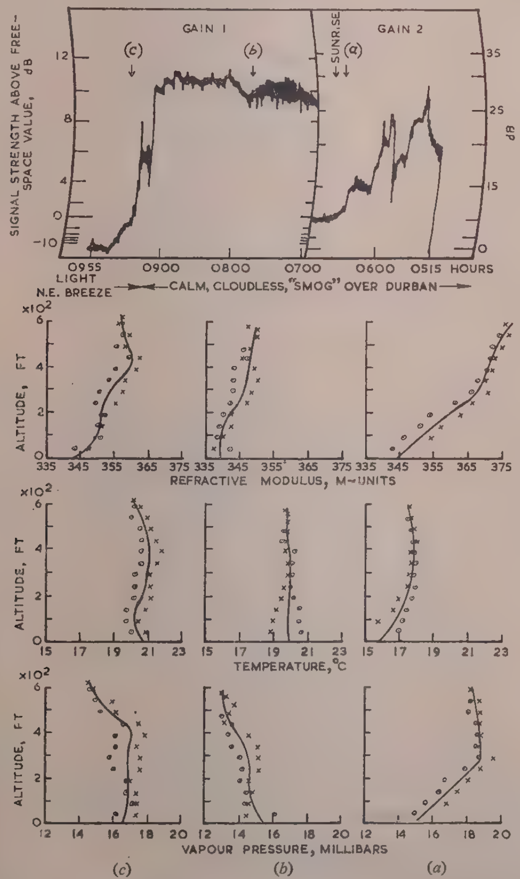


Fig. 3.—Record No. 11; 1st June, 1953; 0515–0955 hours.

× Ascent. ○ Descent.

(a) 0615–0630 hours.
(b) 0735–0805 hours.
(c) 0915–0945 hours.

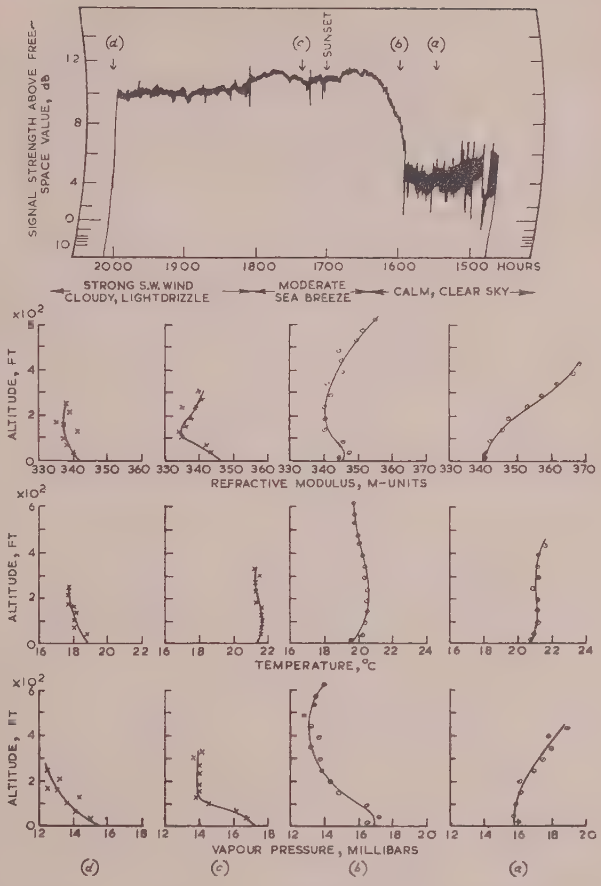


Fig. 4.—Record No. 19; 29th June, 1953; 1445–2005 hours.

× Ascent. ○ Descent.

(a) 1530–1550 hours.
(b) 1600–1615 hours.
(c) 1720–1735 hours
(d) 2015–2030 hours.

and the vertical distribution of atmospheric temperature, vapour pressure, and refractive modulus. Fig. 3 is typical of the results obtained over the sunrise periods, while Fig. 4 shows a rise in signal strength around sunset, and the type of signal obtained under low surface-duct conditions. The record in Fig. 5 was obtained on an occasion when a deep surface duct was present.

(4) QUALITATIVE OBSERVATIONS FROM THE RECORDS

(4.1) Frequency of Duct Formation

Table 1 shows the number of occasions on which ducts of various heights were observed.

There were 25 occasions on which atmospheric soundings gave no indication of the existence of a duct, but on 13 of these occa-

Table 1

INCIDENCE OF DUCTS OF VARIOUS HEIGHTS

Duct height, ft ..	0–50	51–75	76–100	101–125
Number of observations	30	9	5	18
Duct height, ft ..	126–150	151–175	176–200	>200
Number of observations	11	1	6	5

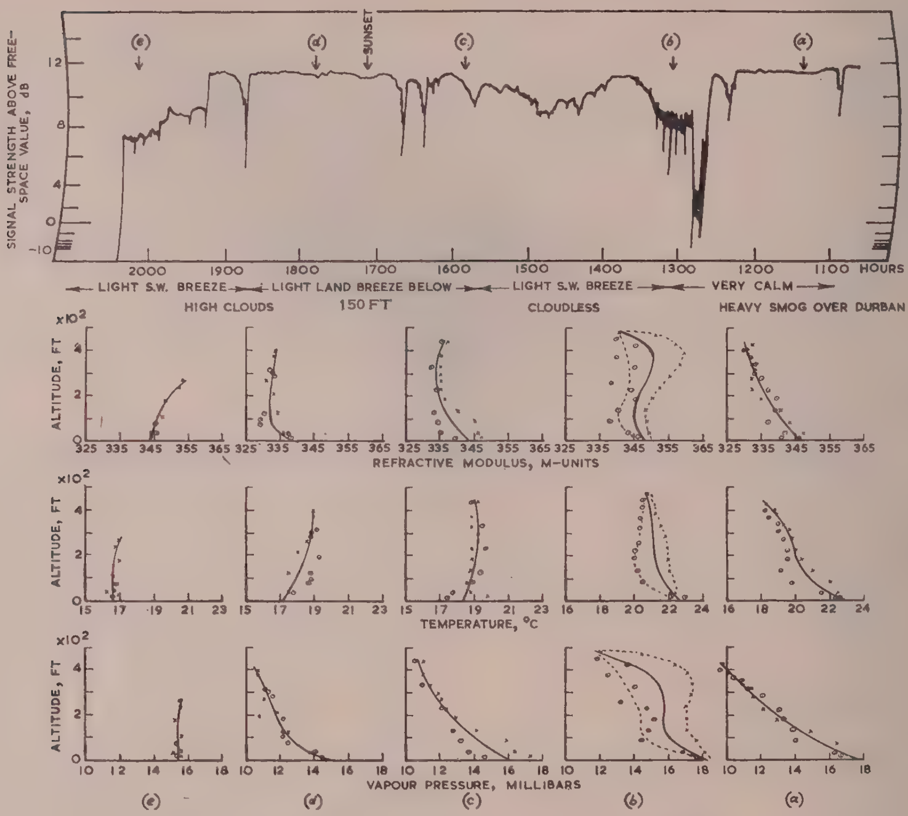


Fig. 5.—Record No. 29; 23rd July, 1953; 1040–2030 hours.

× Ascent. ○ Descent.

(*a*) 1115–1145 hours.
(*b*) 1250–1325 hours.
(*c*) 1540–1610 hours.
(*d*) 1745–1805 hours.
(*e*) 2000–2015 hours.

sions the signal was more than 7·5 dB above the free-space value. Theory indicates that the field strength for a duct height of zero feet would be below the free-space value by an amount of 13 dB for a square-law profile, 18 dB for a fifth-root profile, and 20 dB for a bi-linear profile. The discrepancy may possibly be due to the fact that the atmospheric soundings, taken at a fixed point on the coast, were not sufficiently representative of the conditions over the whole path. If this is assumed to be the case, the magnitude of the signal would provide an indication of the presence of a duct. According to theory, microwaves of 10 cm wavelength are trapped when the duct height exceeds 75 ft. The signal strength for a duct of this height, calculated according to the square-law profile is 1 dB above the free-space value, 2 dB according to the fifth-root profile, and 6 dB according to the bi-linear profile. We may suppose, then, that whenever the signal is more than 6 dB above the free-space value, a duct of minimum height 75 ft exists. Inspection of the records revealed that for 117 hours out of a total operating time of 170 hours the signal exceeded this value, while it was more than 10 dB above the free-space value for a total of 60 hours. This latter figure represents the maximum value of the signal, and would correspond to complete suppression of leakage of the first mode from the top of the duct. The results of the tests therefore indicate that there is a semi-permanent duct of height greater than 75 ft along the Natal coast during the winter months.

(4.2) Effect of Wind on the Duct

The existence of a wind, either from the north-east or the south-west, appears to be favourable to the formation of a duct. Thus, on five occasions the signal strength decreased and the duct disappeared when the wind dropped, whereas on four occasions a rise in signal and formation of a duct accompanied the onset of a wind. An example of this can be seen in Fig. 4. As stated previously, there was sometimes a lag of about one hour between a change in signal and a change in atmospheric conditions, as seen in Fig. 4. On several occasions the reverse effect was noted —on three occasions the onset of a wind coincided with a fall in signal strength, while once a rise in signal occurred within one hour of the disappearance of a breeze.

(4.5) Early-Morning Records

During the winter months a blanket of "smog" forms over Durban in the early hours of the morning. This is an almost daily occurrence, the top of the "smog" blanket appearing to be about 300 ft above sea level. It generally dissipates at about 0900 or 1000 hours. This is believed to indicate the existence of a warm layer of air over a cold damp layer—conditions favourable to the formation of a duct. Ten of the tests covered the early-morning period, and the signal records show certain common characteristics. In general, the signal strength was high before dawn—on two occasions (see Fig. 3) it was more than

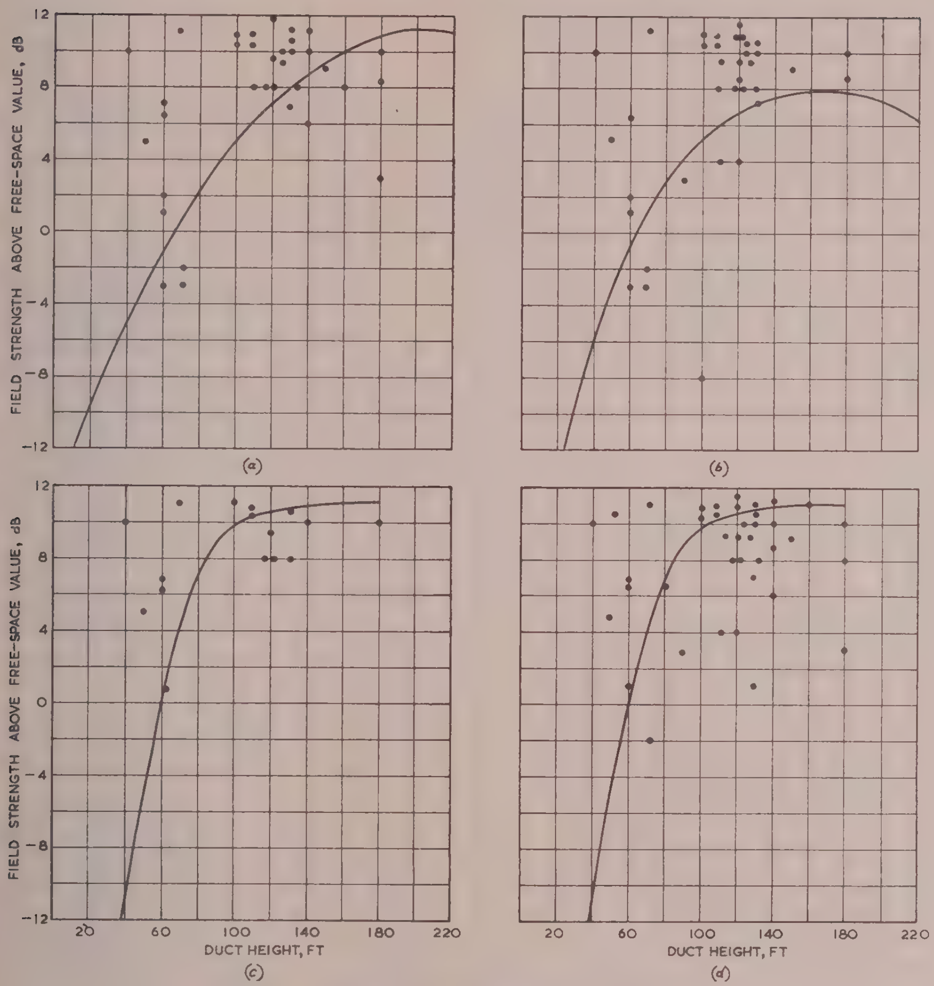


Fig. 6.—Variation of field strength with duct height.

(a) Square-law profile.
(b) Fifth-root profile.
(c) Bi-linear profile; $0{\cdot}85 < |s| < 1{\cdot}15$; $0{\cdot}03 < q < 0{\cdot}04$.
(d) Bi-linear profile; $0{\cdot}7 < |s| < 1{\cdot}3$; $0{\cdot}02 < q < 0{\cdot}05$.

20 dB above the free-space value. However, it generally fell rapidly and assumed a value of 7–10 dB after sunrise. The atmospheric soundings gave no indication of a surface duct, although the possibility of an elevated duct cannot be ruled out. A slight indication of an elevated duct is provided by the soundings shown in Fig. 3. The decrease in signal strength over the sunrise period is accompanied by a change in temperature gradient from positive to negative, and frequently by a change in vapour-pressure gradient from positive to negative. The positive vapour-pressure gradient at low levels militates against the formation of a surface duct. However, it is probable that the gradient becomes negative at higher levels, so that an elevated duct may be formed. On the other hand, the positive gradient may be confined to the coast, where radiative cooling of the land would cause deposition of dew and therefore a decrease in vapour pressure at low levels. This radiative cooling would not occur over the sea, where a surface duct may therefore be formed.

(4.4) Signal Characteristics

During periods of change the signal was subject to rapid fades or scintillations of magnitude 2–4 dB. When the general signal level became constant for long periods, the magnitude of the variations decreased to less than 1 dB. A good example of this is seen in Fig. 4. It was also noted that during periods of rapid change there was an appreciable difference between the results obtained on ascent and on descent of the thermistor unit. When the signal steadied, the results coincided more closely (see Fig. 5).

(4.5) Conditions favouring Formation of a Duct

Ducts are formed under conditions of a positive temperature gradient and a high negative gradient of vapour pressure. Examination of the records showed that in only 9 out of 56 soundings when a surface duct was present was the temperature gradient positive. It is evident that the principal cause of duct formation in this area is a high lapse rate of vapour pressure.

(5) QUANTITATIVE OBSERVATIONS FROM THE RECORDS

Correlation between measured signal strengths and the values calculated according to the various modified index profiles, using only the first mode, was investigated as follows. The average signal strength during the time of a sounding was determined from the records and plotted against duct height. On the same graph was plotted the theoretical variation of signal strength

with duct height. Only those results were plotted for which the measured profile was an acceptable fit to the theoretical profile. The goodness of fit was estimated visually by superimposing a tracing of the theoretical profile over the actual one.

The results for the square-law profile are given in Fig. 6(*a*). It is seen that there is quite a fair degree of correlation, both as regards the absolute value of signal strength and as regards the variation with duct height. The spread of the results appears to be considerable, but it can be seen that an error of only 2 dB in signal strength and 15 ft in duct height would account for most of the discrepancies between measured and calculated values.

Fig. 6(*b*) shows the results for the fifth-root profile. Again correlation is quite fair. This profile does not fit the actual profiles very well, particularly at altitudes below about 50 ft, where the *M*-gradient is much greater than it is found to be in practice.

Theoretical values for the bi-linear profile were only available for profiles such that the ratio, s, between the gradients above and below the cut minimum was equal to -1, the gradient above the duct minimum, q, being standard, i.e. $0\cdot036$ *M*-units per foot. In Fig. 6(*c*) are plotted those observed values for which the ratio $|s|$ lies within the limits $0\cdot85$ and $1\cdot15$, and the gradient q in the range $0\cdot03$–$0\cdot04$. In Fig. 6(*d*) are plotted those observed values which lie within the extended ranges $0\cdot7 < |s| < 1\cdot3$, $0\cdot02 < q < 0\cdot05$.

A number of the soundings gave substantially linear *M*-profiles. Variation of measured signal strength with positive gradient of the *M*-profile in these cases is illustrated in Fig. 7, together with

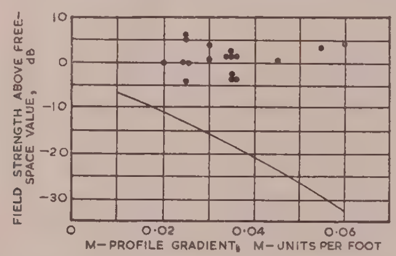


Fig. 7.—Variation of field strength with gradient of linear profile.

the theoretical variation. Only those observations when the signal strength was low have been plotted. Occasions when a high signal strength was obtained with a linear positive *M*-gradient have been omitted, as it was suspected that some other effect—possibly an elevated duct—was in operation at these times. Correlation is not very good, and there is a large discrepancy between the measured and the calculated values—about 16 dB—the observed values being higher than the theoretical. Similar results were obtained in propagation experiments across Massachusetts Bay in 1942–44.[1] Here signal strengths on a 9 cm link over a non-optical sea path 41 miles long were, on an average, about 10 dB above the values calculated on the basis of the bi-linear model, for duct heights less than about 30 ft. Again, Swarm, Ghose and Keitel[5] found, for a 150-mile path over land, that the observed signal strength on frequencies of 100 and 500 Mc/s was, on the average, $11\cdot5$ dB above the value calculated for a linear modified-index profile.

(6) STATISTICAL ANALYSIS

A statistical analysis was undertaken to investigate whether any diurnal variation in signal strength was detectable. The mean hourly signal values during the operating periods were averaged for the whole series, and are shown plotted in Fig. 8. It is seen that there is a fall in signal strength from before dawn to about mid-morning, after which there is a slight rise reaching

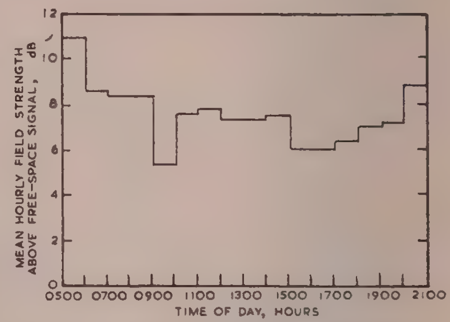


Fig. 8.—Diurnal variation of signal strength.

a peak shortly after midday. Thereafter the signal strength falls again, but rises about sunset. This average variation was plainly discernible in several of the records. An attempt to correlate these variations with the incidence of ducts was not very successful, owing to the paucity of soundings in the early-morning and evening periods. A possible explanation of this diurnal variation in signal strength is as follows: The sharp fall over the sunrise period is due to the disappearance of the land breeze, which gives rise to a temperature inversion and a high lapse rate of humidity over the sea. The rise at mid-morning may be associated with the onset of the prevailing winds, which, as noted previously, is frequently accompanied by a rise in signal strength. This generally diminishes in force towards sunset, and the signal strength falls. After sunset there is often a rise in signal strength, which could be accounted for by the onset of a land breeze.

(7) CONCLUSIONS

These experiments show that during the winter months a duct of height approximately 120 ft exists for a considerable portion of the time along the Natal coast. The signal strength at a height of 47 ft in the duct is 6–10 dB above the free-space value. The absolute value of the signal strength, and its variation with duct height, are given moderately well by the mode theory, using only the first mode, and using either a square-law, fifth-root or bi-linear profile.

(8) ACKNOWLEDGMENTS

The investigation was carried out under the auspices of the Department of Electrical Engineering, University of Natal. The equipment was constructed and calibrated partly in the laboratories of the Department, and partly in the workshops and laboratories of the Natal Technical College.

(9) REFERENCES

(1) Kerr, D. H.: "Propagation of Short Radio Waves" (McGraw-Hill, 1951).
(2) Booker, H. G., and Walkinshaw, W.: "The Mode Theory of Tropospheric Refraction and its Relation to Wave-Guides and Diffraction," *Meteorological Factors in Radio-Wave Propagation* (The Physical Society, 1946), p. 80.
(3) Megaw, E. C. S.: "Experimental Studies of the Propagation of Very Short Radio Waves," *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 79.
(4) Norton, K. A.: "The Calculation of Ground-Wave Field Intensity over a Finitely Conducting Spherical Earth," *Proceedings of The Institute of Radio Engineers*, 1941, **29**, p. 623.
(5) Swarm, H. M., Ghose, R. N., and Keitel, G. H.: "Tropospheric Propagation in Horizontally Stratified Media over Rough Terrain," Convention Record of the Institute of Radio Engineers, Part 2, 1953 National Convention, p. 77.

621.396.11

Paper No. 1783 R
Mar. 1955

THE LAUNCHING OF A PLANE SURFACE WAVE

By G. J. RICH, M.Sc.

(*The paper was first received* 14*th July, and in revised form* 20*th October*, 1954.)

SUMMARY

The subject of the paper is a theoretical and experimental investigation of surface-wave propagation over a plane conductor with dielectric coating. Special reference is made to the problem of launching such a wave.

Two experimental surfaces have been constructed, and a series of measurements has confirmed the properties of the waves as reported in the literature. A number of trials of various possible launching devices have been conducted, and eventually a launcher of moderate efficiency was found. A theoretical investigation of the launching efficiency of an aperture containing a surface-wave field of the correct form was conducted, and the launcher already mentioned was used to generate such a field in an aperture. Measurements were taken to find the launching efficiency of the aperture, and the results were found to be in agreement with the theory for small apertures. The deviations from the theoretical law for larger apertures are attributed to the fact that the aperture field differed slightly from the theoretical ideal in these cases.

LIST OF PRINCIPAL SYMBOLS

a = Thickness of dielectric coating.
a = Suffix denoting location within the aperture.
c = Velocity of light *in vacuo*.
C_0 = Cosine of Brewster angle.
E = Electric field strength.
$F(u)$ = Fresnel integral.
h = Height of aperture.
H = Magnetic field strength.
J_z = Surface current density.
k = Propagation coefficient of surface-wave $= 2\pi/\lambda$.
k_2, k_3 = Space constants for dielectric and outer media.
n = Mode number.
p_n = Solution of transcendental equation determining characteristics of surface wave.
P = Power.
r = Distance beyond diffracting aperture.
r = Suffix denoting location distant r from the aperture.
$S_0' = S_0 + jX$ = Sine of Brewster angle.
u, jw = Arguments of Fresnel integral.
v = Phase velocity of surface wave.
x, y, z = Co-ordinates.
z = Argument of error function.
$\gamma_z = \alpha_z + j\beta_z$ = Propagation coefficient of surface wave in z-direction.
ϵ = Permittivity.
η = Launching efficiency.
ζ = Variable of integration.
θ_0 = Brewster angle.
λ = Free-space wavelength of radiation.
λ_{cn} = Cut-off wavelength for nth mode.
λ_z = Phase wavelength of surface wave.
μ = Permeability.
ρ = Ratio of magnetic to electric field.
σ = Conductivity.
τ = Variable of integration.
ω = Angular frequency.
$\Delta y_1, \Delta y_2, \Delta y_3$ = Differences between various values of y.
0, 2, 3 = Suffixes denoting various media (0—free space, 2—dielectric, 3—free space).

(1) INTRODUCTION

Since Sommerfeld[1] and Zenneck[2] first formulated the equations of the electromagnetic surface wave as a solution of Maxwell's equations for certain boundary conditions half a century ago, the possibility and practicability of launching such waves has attracted much attention. The question of the existence of the surface wave in the radiation over the earth from a horizontal dipole has been the subject of a controversy too involved and protracted to be described briefly: references[3–20] are given to the most important papers on the subject. However, the excitation of surface waves in the laboratory is a more practical proposition. The conical launchers used by Goubau[21] for the single-wire transmission line are well known, and analagous devices have been used to excite plane surface waves by Rotman,[22] Reynolds and Lucke,[23] and Erlich and Newkirk.[24] Interest has centred mainly upon the use of plane guiding surfaces as microwave end-fire aerials. A survey of the various possible types of surface wave has been written by Barlow and Cullen.[25]

The wave described by Zenneck is identical with that incident upon a plane conductor at the Brewster angle, where no reflection occurs. The derivation is carried out by Stratton[26] for all possible guiding surfaces. For the case of an imperfect conductor, the wave is exponentially attenuated in the direction of propagation, and the planes of constant phase are inclined to the guiding surface at an angle

$$\psi_0 = \arc\tan \sqrt{\left(\frac{2\mu_2\sigma_1}{\mu_1\epsilon_2\omega}\right)} \quad . \quad . \quad . \quad . \quad (1)$$

The suffixes 1 and 2 refer to the conductor and the dielectric (air) respectively. Owing to the finite conductivity, there is a component of the electric vector in the direction of propagation. The wave is predominantly of the transverse magnetic type. The intensities of all the field components fall off according to an exponential law

$$E, H \propto \varepsilon^{jkS_0'y} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where y is measured normally upwards from the surface.

When the conducting surface is "loaded" either by having corrugations cut in it or being coated with a dielectric, the Brewster angle becomes complex. In the case of a perfect conductor coated with a perfect dielectric, the Brewster angle is purely imaginary. This view of the phenomenon regards the conductor and its coating as together constituting a surface whose electrical constants are complex quantities.

A more rigorous analysis may be carried out by solving Maxwell's equations using the conditions at the two interfaces (dielectric-air and conductor-dielectric) as boundary conditions. This is the procedure followed by Attwood,[27] who deduces the possible existence of two types of wave under these conditions.

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Mr. Rich is at the Radar Research Establishment.

One of these radiates away from the surface; the other is bound to it and is the subject of the present paper. Attwood assumes a TM mode and deduces the following relations for the field components inside the dielectric:

$$E_{z2} = jJ_z \frac{k_2}{\omega\epsilon_2} \sin(k_2 y) \quad . \quad . \quad . \quad . \quad . \quad (3)$$

$$E_{y2} = J_z \frac{\beta}{\omega\epsilon_2} \cos(k_2 y) \quad . \quad . \quad . \quad . \quad . \quad (4)$$

$$H_{x2} = -J_z \cos(k_2 y) \quad . \quad . \quad . \quad . \quad . \quad (5)$$

and for the field outside the dielectric:

$$E_{z3} = jJ_z \frac{k_2}{\omega\epsilon_2} \varepsilon^{k_3 a} \sin(k_2 a)\varepsilon^{-k_3 y} \quad . \quad . \quad . \quad . \quad (6)$$

$$E_{y3} = J_z \frac{k_2}{k_3} \frac{\beta}{\omega\epsilon_2} \varepsilon^{k_3 a} \sin(k_2 a)\varepsilon^{-k_3 y} \quad . \quad . \quad . \quad (7)$$

$$H_{x3} = -J_z \frac{k_2}{k_3} \frac{\epsilon_3}{\epsilon_2} \varepsilon^{k_3 a} \sin(k_2 u)\varepsilon^{-k_3 y} \quad . \quad . \quad . \quad (8)$$

The conductivity of the metal is assumed to be infinite; the suffixes 2 and 3 refer to the dielectric and the air, respectively; J_z is the surface current on the conductor per unit depth in the x-direction. Propagation is in the z-direction, and y is perpendicular to the surface. The term $\exp[j(\omega t - \beta z)]$ has been omitted. The relations

$$k_2^2 = \omega^2\mu_2\epsilon_2 - \beta^2 \quad . \quad . \quad . \quad . \quad . \quad (9)$$

$$k_3^2 = \beta^2 - \omega^2\mu_3\epsilon_3 \quad . \quad . \quad . \quad . \quad . \quad (10)$$

and $$k_2^2\left[1 + \frac{\epsilon_3^2}{\epsilon_2^2}\tan^2(k_2 a)\right] = \omega^2(\mu_2\epsilon_2 - \mu_3\epsilon_3) \quad . \quad . \quad (11)$$

determine the constants. Eqn. (11) permits of more than one solution with suitable values of ϵ_2, ω and a, so that more than one mode can be set up under suitable conditions. But there is no cut-off for the lowest mode because eqn. (11) always has at least one solution. It can be seen from eqns. (3)–(8) that the field has sinusoidal distribution inside the dielectric and dies away exponentially outside it. Also it can be deduced that increasing the value of a enables the higher modes to propagate and that in this case the power is concentrated nearer to the surface. A larger proportion of the power is then carried within the dielectric. Figs. 1(a) and 1(c) show the field distributions and the lines of electric force for this mode.

Shersbie-Harvey[28] quotes field equations similar to those of Attwood using a rather different notation. His constant determining the modes is given by

$$\tan\left[2\pi p_n\sqrt{\left(\frac{\mu\epsilon}{\mu_0\epsilon_0} - 1\right)}\frac{a}{\lambda}\right] = \frac{\epsilon}{\epsilon_0}\sqrt{\left(\frac{1}{p_n^2} - 1\right)} \quad . \quad (12)$$

p_n is the $(n+1)$th solution of this equation. The other symbols are similar to Attwood's μ_2, ϵ_2, μ_3, ϵ_3 with the zero suffixes corresponding to Attwood's 3's. Eqns. (11), (12) can be shown to be identical by putting

$$k_2 = \frac{2\pi}{\lambda}\sqrt{\left(\frac{\mu\epsilon}{\mu_0\epsilon_0} - 1\right)}p_n \quad . \quad . \quad . \quad . \quad (13)$$

Shersbie-Harvey also describes transverse electric modes. These are very similar to the TM modes, but p_n is in this case a solution of

$$\cot\left[2\pi p_n\sqrt{\left(\frac{\mu\epsilon}{\mu_0\epsilon_0} - 1\right)}\frac{a}{\lambda}\right] = -\frac{\mu}{\mu_0}\sqrt{\left(\frac{1}{p_n^2} - 1\right)} \quad . \quad (14)$$

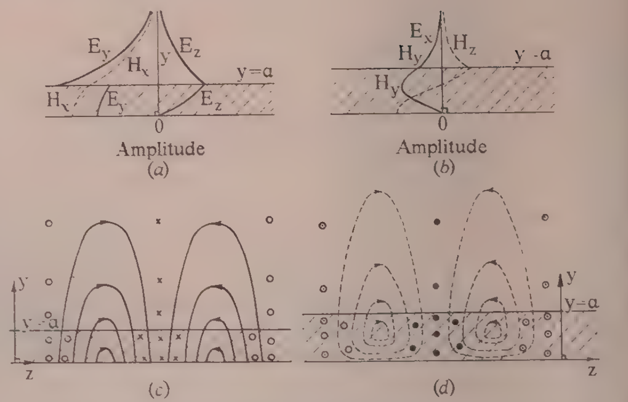


Fig. 1.—Field components and field lines.

(a) Field components; TM_0 mode.
(b) Field components; TE_0 mode.
(c) Field lines; TM_0 mode.
(d) Field lines; TE_0 mode.
———, ○, ● Electric-field lines.
- - - -, ○, × Magnetic-field lines.

Eqn. (14) has no real solutions if a is sufficiently small; so the lowest mode in this case has a cut-off wavelength. This is given by

$$\lambda_{C0} = 4a\sqrt{\left(\frac{\mu\epsilon}{\mu_0\epsilon_0} - 1\right)} \quad . \quad . \quad . \quad . \quad (15)$$

A map of the field lines for the TE_0 mode is shown in Fig. 1(d).

The launching of a Sommerfeld–Goubau wave upon a wire is facilitated because the radial variation of the field follows a Hankel-function distribution which approximates to the inverse radius law obeyed by fields inside a coaxial line. The wave is therefore usually launched by flaring the outer conductor of a coaxial line into a cone and continuing the inner conductor to form the transmission line. However, the exponential decay of the fields above a coated conductor does not approximate at all closely to the constant or sinusoidal field distribution inside waveguides and strip transmission lines, and the launching of a pure surface wave presents some difficulty.

Macfarlane[29] deduces that a Zenneck wave, if it is to travel an infinite distance over an imperfectly conducting earth, must be launched from an aperture of infinite height containing a field of the form given in eqn. (2). For a finite launching aperture, the range of the wave is restricted in this case to a distance $h \operatorname{cosec} \theta_0$, where h is the height of the aperture and θ_0 is the Brewster angle. Beyond this distance there exists only the diffraction field of the aperture. For the lossless dielectric-coated surface θ_0 becomes purely imaginary, and Macfarlane shows that even for a finite aperture the range of the wave is infinite. However, not all the energy goes into the surface wave, and it will be shown later how the height of the aperture determines the proportion of the total energy that goes into the surface wave.

The launching devices so far invented have been various in form. In those experiments which have dealt principally with the surface wave as a means of making a compact aerial, the emphasis has not been strongly placed upon getting a high proportion of the available power into the surface wave. The experimenters have aimed chiefly at making the surface wave interact with the radiated field to produce a shaped beam. It is easy to launch a surface wave which will, at a distance, become predominant over the radiation field near the surface, but it is not easy to concentrate nearly all the power into the surface wave.

(2) OUTLINE OF THE INVESTIGATIONS

The object of the investigations was to confirm the properties of this surface wave that are reported in the literature and then to find a launcher that would convert most of its input power into a surface-wave mode.

A dielectric surface was first constructed, and an automatic recording mechanism for the electric field was set up. Various launching devices were tried, and it was found to be easy to launch a wave having approximately the exponential field distribution of the surface wave. However, efficient launching was not easily achieved, and after numerous attempts a launcher of about 50% efficiency was devised. Launching efficiency is defined as the ratio of power in the surface wave to total power. A theoretical expression for the launching efficiency of an aperture like that described by Macfarlane[29] was deduced, and the launcher mentioned above was used to set up a field of the required configuration in an aperture whose height could be varied. The experimental results were compared with the theoretical deductions and they were found to agree closely for small apertures (less than one wavelength). The theory predicts that efficiencies of the order of 90% can be obtained from apertures of about 2λ, but it seems that the field in the launching aperture must approximate very closely to the theoretical one.

It is hoped in the future to investigate and ultimately to develop a launcher of high efficiency and also to investigate the effect of modifying the dielectric to launch the wave from the surface into space.

The experiments were conducted at a wavelength of about 3cm (X-band). Polystyrene with a nominal thickness of $\frac{1}{16}$in was used as the coating medium on a brass base. Its dielectric constant is about 2·5; an accurate determination of this is not essential to the success of the measurements. A thicker dielectric coat would make the field cling inconveniently closely to the surface, thus rendering more difficult accurate measurements of the field above the dielectric. To use a shorter wavelength, say 8mm, would have made the scale of the apparatus much smaller, but attenuation in the dielectric would increase, and a thinner dielectric would have had to be found. Reducing the apparatus to this small scale would have made it much more difficult to construct; also $\frac{1}{16}$in polystyrene will just support a TE_0 mode at 8mm wavelength. With $\lambda = 3$cm and $a = \frac{1}{16}$in, only the TM_0 mode will propagate.

(3) GENERAL DESCRIPTION OF THE APPARATUS

Two guiding surfaces have been constructed so far in this investigation. The first was 6in wide and 4ft long, later extended to 6ft. The second surface is 6ft × 1ft, and like the first it is coated with $\frac{1}{16}$in polystyrene sheet, which is in both cases stuck down at the edges with small pieces of p.v.c. adhesive tape.

The change from the 6in to the 1ft surface was made because it was found that a wider aperture was needed to make the field approximate more closely to the theoretical form. The launcher was placed at one end of the surface so that the wave was propagated longitudinally; at the further end the surface was terminated either by a load consisting of wedges of resistive card projecting from slots in a wooden block or by a short-circuit consisting simply of a length of 3in × 1½in rectangular waveguide placed at right-angles across the surface.

Radio-frequency power at a wavelength of 3cm was supplied from a type CV129 klystron. The output was modulated at 3kc/s, a modulated power supply being used to supply the oscillator. A crystal detector mounted in waveguide and a 3kc/s tuned amplifier were used for the detection of the radiation. A sensitive galvanometer was used for the measurement of standing waves in the input waveguide since there was enough power available in this case to operate the instrument. Continuous-wave power was then used. Fig. 2 gives a schematic of the apparatus.

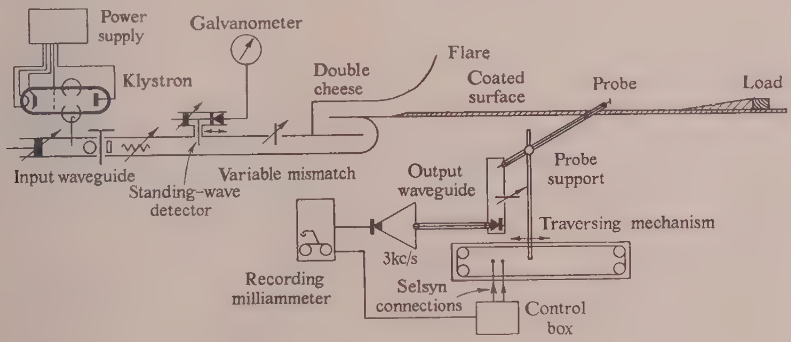


Fig. 2.—Schematic of the apparatus.

One other major permanent item of the apparatus is the traversing mechanism for the probe. This consists of an electrically-driven carriage moving along a track 3ft in length. The movement of the carriage is transmitted by means of reduction gearing and a Selsyn connection to the recording milliammeter. The gearing is so adjusted that the carriage and the recorder paper move equal distances. The output of the amplifier is brought directly to the milliammeter terminals. By means of this automatic recording system a large quantity of data can be gathered in a very short time.

The probe and its ancillary waveguides are arranged to slide on a vertical steel rod fixed in the clamps of the travelling carriage. The waveguide, to which the probe is rightly attached, is clamped to the rod when the mechanism is in operation. There is a height adjustment, consisting of a spring and threaded rod, which is used to preset the probe to a given position.

The probe is a half-wave dipole fitted on the end of a coaxial line made by inserting a length of insulated wire down the centre of a brass tube. The inner conductor is brought out to form one branch of the dipole; the other is soldered to the tube, which has an 8mm slot cut in its end to act as a choke. The output from the probe is brought to a coaxial socket which screws on to a coaxial-to-waveguide transformer forming the top item in the waveguide run. This system has the advantage of lightness and rigidity, but the probe is unsatisfactory in some respects since it distorts the field when placed in certain positions. However, attempts to improve it were unsuccessful, and this design was adopted only after experiments with less satisfactory probes.

The crystal holder can be seen at the lower end of the output waveguide run in Fig. 2. Between it and the input transformer is inserted a mismatch unit consisting of a travelling screw. This is used to correct as far as possible the various mismatches existing in the probe and waveguides. It is always tuned for maximum crystal output.

(4) DEVELOPMENT OF A LAUNCHING DEVICE

The next phase of the experiments was to try to make an efficient launcher. Launching efficiency was measured by the following process. The launching device under test was first of all matched to the surface by terminating the latter with the resistive-card load and adjusting the mismatch unit shown in the input waveguide run in Fig. 2 until the standing-wave meter on the right of the mismatch unit showed that there was a good match (the galvanometer was generally used here). The surface-wave load had a voltage standing-wave ratio of about 0·97. The short-circuit was then placed across the surface, and the voltage standing-wave ratio was again measured. The launching efficiency was then equal to the voltage reflection coefficient. This can be shown as follows: Suppose that the launching efficiency is η, and P is the total input power, the power in the surface wave will be ηP. This will be reflected by the short-circuit, but the short-circuit is not a serious obstacle to the rest of the power, which will radiate away into space. So the power arriving back at the horn is ηP, and $\eta^2 P$ will go back into the waveguide, since by the reciprocity theorem the launching and collecting efficiencies must be equal. The power reflection coefficient is therefore η^2, and the voltage reflection coefficient is η.

Numerous unsuccessful launching devices were tested, but only one will be mentioned here. In view of Dr. Cullen's theoretical investigations,[30] slot sources were constructed. These, according to the theory, should radiate a surface wave with high efficiency when placed at a certain critical height above the plane. However, no success was achieved either with a long slot extending right across the plane or with a short slot and parabolic reflector. In no case did the measured launching efficiency exceed 20%.

The first launcher of moderate efficiency which was tested consisted of a sectoral horn with lens correction having an aluminium flare fitted to the top of the aperture. The flare was bent to the shape that gave the highest efficiency. A number of measurements led to the following conclusions:

(*a*) A curved flare provides greater launching efficiency than a straight one.

(*b*) The best results are obtained when the flare rises smoothly from the top surface of the horn and its curvature increases with the distance from the horn.

The highest efficiency recorded with this device was 42%. Up to this point all measurements had been conducted on the 6 in surface, but in order to minimize edge effects and to try to get a flatter phase front at the launcher, the width was now increased to 12 in.

The form of launcher provisionally adopted at this stage for further measurements was a modification of the sectoral horn and flare that had been the best launcher used with the 6 in surface. The flare remained, but the horn was replaced by a "double cheese" because the horn necessary to give a really flat phase front 12 in across would have been inconveniently long, even with lens correction.

The double cheese is only 3 in deep for a 12 in aperture, if the feed is in the aperture plane. The feed was an open-ended waveguide situated in the lower half of the cheese. Double cheeses are described by Kiely, Collins and Evans.[31] This particular cheese is illustrated in Fig. 3. The end of the polystyrene coating is filed down to a sharp edge and comes nearly to the mouth of the cheese itself. The brass sheet forming the conducting part of the surface is the centre tongue of the cheese, being cut at the end to the required contour.

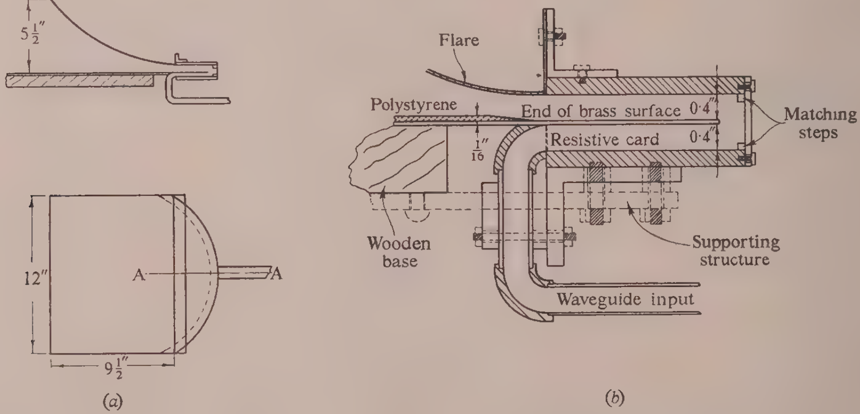


Fig. 3.—Double-cheese launcher.

(*a*) Complete launcher.
(*b*) Detail of double cheese.

The efficiency of this launcher when measured was 50%. Attempts were made to improve upon this figure but with little success. This launcher, although its efficiency was still far below the figure desired (80% or 90%), was considered to be good enough to be used to generate a surface wave for the next stage in the experiments.

(5) THEORETICAL TREATMENT OF A FINITE APERTURE

The methods used in this Section are based on Macfarlane's paper[29]; he in turn uses the methods and notation of Booker and Clemmow.[32] This notation differs from the one adopted here in assuming that propagation is parallel to the x-axis. This has been changed to the z-axis to bring the co-ordinates into line with those hitherto employed.

It is assumed that an aperture of height h contains fields of "curtailed exponential" form:

$$E_y = C_0 \varepsilon^{jkS_0' y} \quad . \quad . \quad . \quad (16)$$

$$0 \leqslant y \leqslant h$$

$$H_x = \rho \varepsilon^{jkS_0' y} \quad . \quad . \quad . \quad (17)$$

S_0' and C_0 are respectively the sine and cosine of the Brewster angle. In this case as the Brewster angle is purely imaginary, we can put

$$S_0' = jX, \; C_0 = \sqrt{(1 + X^2)}$$

Macfarlane goes on to show that the field distribution beyond the aperture consists of both a surface wave and a diffraction field. The complete expression for the electric field beyond the aperture is

$$E_y = \frac{1}{\sqrt{2}} \varepsilon^{-jk\sqrt{(1+X^2)}z} \varepsilon^{-kXy} \varepsilon^{j\pi/4} \sqrt{(1 + X^2)}$$

$$\left\{ F\left[\left(\frac{y-h}{r} + jX\right)\sqrt{\left(\frac{\pi r}{\lambda}\right)}\right] - \varepsilon^{-2kXh} F\left[\left(\frac{y+h}{r} + jX\right)\sqrt{\left(\frac{\pi r}{\lambda}\right)}\right]\right\} \quad . \quad (18)$$

where r is the distance beyond the aperture, and

$$F\left[\left(\frac{y-h}{r} + jX\right)\sqrt{\left(\frac{\pi r}{\lambda}\right)}\right], \; F\left[\left(\frac{y+h}{r} + jX\right)\sqrt{\left(\frac{\pi r}{\lambda}\right)}\right]$$

are Fresnel integrals defined by

$$F(u) = \sqrt{\left(\frac{2}{\pi}\right)} \int_u^\infty \varepsilon^{-j\zeta^2} d\zeta \quad . \quad . \quad . \quad . \quad (19)$$

When r is sufficiently large, the arguments of both Fresnel integrals become purely imaginary and equal to $jX\sqrt{(\pi r/\lambda)}$. The Fresnel integrals may then be replaced by the following asymptotic expansion:

$$F(jw) = \sqrt{(2)} \varepsilon^{-j\pi/4} - \frac{\varepsilon^{jw^2}}{\sqrt{(2\pi)}w} \quad . \quad . \quad . \quad (20)$$

of which only the first term need be taken.

Eqn. (18) now becomes

$$E_y = \varepsilon^{-jk\sqrt{(1+X^2)}z} \varepsilon^{-kXy} \sqrt{(1 + X^2)}(1 - \varepsilon^{-2kXh}) \quad . \quad (21)$$

This is the field a long distance from the aperture and includes only the surface-wave component. The launching efficiency may therefore be computed by comparing the power flow at this distance with that through the aperture. The power flow is given by

$$P_r = \mathscr{R}\left(\int_0^\infty E_y H_x^* dy\right)$$

for this case, where H^* denotes the complex conjugate of H.

$$H_x = \rho \varepsilon^{-jk\sqrt{(1+X^2)}z} \varepsilon^{-kXy}(1 - \varepsilon^{-2kXh}) \quad . \quad . \quad (22)$$

$$H_x^* = \rho \varepsilon^{jk\sqrt{(1+X^2)}z} \varepsilon^{-kXy}(1 - \varepsilon^{-2kXH})$$

therefore $P_r = \rho \int_0^\infty (1 - \varepsilon^{-2kXh})^2 \sqrt{(1 + X^2)} \varepsilon^{-2kXy} dy$

$$= \rho \frac{\sqrt{(1 + X^2)}}{2kX}(1 - \varepsilon^{-2kXh})^2 \quad . \quad . \quad . \quad . \quad (23)$$

The power flow through the aperture is given by

$$P_a = \rho \int_0^h \sqrt{(1 + X^2)} \varepsilon^{-2kXy} dy$$

$$= \rho \frac{\sqrt{(1 + X^2)}}{2kX}(1 - \varepsilon^{-2kXh}) \quad . \quad . \quad . \quad (24)$$

The launching efficiency is therefore

$$\eta = \frac{P_r}{P_a} = 1 - \varepsilon^{-2kXh} \quad . \quad . \quad . \quad . \quad (25)$$

which is a very simple result.

Before this result can be tested, it is necessary to find a value for X. X is related both to the exponential decrement of the field components and to the phase velocity of the surface wave. Comparing Shersbie-Harvey's formula

$$E_y \propto \exp\left\{-\left[\frac{2\pi}{\lambda}\sqrt{(1 - p_n^2)}\sqrt{\left(\frac{\mu\epsilon}{\mu_0\epsilon_0} - 1\right)}\right]y\right\} \quad . \quad (26)$$

with Macfarlane's $\qquad E_y \propto \varepsilon^{-kXy}$

it is seen that $\qquad X = \sqrt{(1 - p_n^2)}\sqrt{\left(\frac{\mu\epsilon}{\mu_0\epsilon_0} - 1\right)} \quad . \quad . \quad . \quad (27)$

X can therefore be computed from a solution of the transcendental eqn. (12).

If v is the phase velocity of the surface wave and λ_z its phase-wavelength, $v/c = \lambda_z/\lambda$. Attwood gives

$$E_y \propto \varepsilon^{-k_3 y}$$

Comparing this with Macfarlane's expression and substituting in eqn. (10) we have

$$\frac{4\pi^2}{\lambda^2}X^2 = \beta^2 - \omega^2\mu_3\epsilon_3 \quad . \quad . \quad . \quad . \quad (28)$$

β is the phase-change coefficient in the z-direction and is equal to $2\pi/\lambda_z$. Substituting $\omega = 2\pi c/\lambda$, and remembering that $c = 1/\sqrt{(\mu_3\epsilon_3)}$, we have

$$\frac{\lambda_z}{\lambda} = \frac{1}{\sqrt{(1 + X^2)}} \quad . \quad . \quad . \quad . \quad . \quad (29)$$

or

$$v = \frac{c}{\sqrt{(1 + X^2)}} \quad . \quad . \quad . \quad . \quad . \quad (30)$$

This is similar to the case of the Zenneck wave over an imperfect conductor, where $v = c \sec \theta_0$.

X can therefore be evaluated either by computation from eqns. (12) and (27) or by measuring the phase velocity of the surface wave and using eqn. (30). In the next Section it will be described how both these methods were employed to find X and how the results compared.

(6) EXPERIMENTAL TEST OF THE THEORY

Before the theoretical curve for launching efficiency

$$\text{Efficiency} = 1 - \varepsilon^{-2kXh}$$

can be drawn, a value must be found for the parameter X. This can be done either by measuring λ_z/λ or by computing a value for X from the field equations. The computation was carried out from the equations given by Shersbie-Harvey. The first

step was to solve eqn. (12) to find a value for p_0. Newton's method was used, by putting

$$f(p_0) \equiv \tan\left[\frac{2\pi}{\lambda}\sqrt{\left(\frac{\mu\epsilon}{\mu_0\epsilon_0}-1\right)}ap_0\right] - \frac{\epsilon}{\epsilon_0}\sqrt{\left(\frac{1}{p_0^2}-1\right)} = 0 \quad . \quad (31)$$

In the case in question $\mu/\mu_0 = 1$, $\epsilon/\epsilon_0 = 2\cdot5$,* $a = 0\cdot17$ cm.† and $\lambda = 3\cdot15$ cm. Eqn. (31) then becomes

$$f(p_0) \equiv \tan(0\cdot4154\,p_0) - 2\cdot5\sqrt{\left(\frac{1}{p_0^2}-1\right)} = 0$$

$$\text{and} \quad f'(p_0) = 0\cdot4154\sec^2(0\cdot4154\,p_0) + \frac{2\cdot5}{p_0^3}\left(\frac{1}{p_0^2}-1\right)^{-\frac{1}{2}} \quad . \quad (32)$$

Taking $p_0 = 0\cdot99$ as a first approximation p_0 is soon found to be $0\cdot9853$, to an accuracy of one part in 10^4.

Substituting into eqn. (27), we have

$$X = 0\cdot2111.$$

The practical measurement of X was undertaken as follows: A short-circuit was placed on the end of the surface so that a large standing-wave was produced. With the amplifier gain at maximum setting, the minima were seen to be very sharp. Using manual traverse on the probe mechanism, the probe was set to a minimum, using a bracketing method. The probe was then moved along the median line of the surface, and 55 minima were counted in the length of its traverse. The position of the 56th minimum was determined in the same way as that of the first. The intervening distance was measured with a steel rule, and was estimated at $84\cdot925$ cm. So

$$55/2 \times \lambda_z = 84\cdot925\,\text{cm}.$$

This leads to a figure of $3\cdot088$ cm $\pm\,0\cdot125\%$ for λ_z, allowing for errors in the steel rule and the measurements. λ was measured with a wavemeter, newly-calibrated against a quartz crystal and correct to 1 part in 10^4, and a value of $3\cdot1553$ cm $\pm\,0\cdot01\%$ was obtained.

Therefore $\quad v/c = \lambda_z/\lambda = 0\cdot9787 \pm 0\cdot135\%$

so, from eqn. (30), $X = \sqrt{\left(\dfrac{c^2}{v^2}-1\right)} = 0\cdot210 \pm 3\%$

Owing to the form of the relation between v/c and X, all errors are considerably multiplied when X is calculated, but even so this value agrees closely with the theoretical one. This may be taken as a check on the correctness of the various field equations.

In Fig. 4 is seen the theoretical curve for the launching efficiency. It can be seen that, at first, the launching efficiency increases very rapidly with increasing aperture, but that from about $h = 5$ cm the curve flattens out and then approaches 100% asymptotically. It seems from this curve that, in theory, launching efficiencies of up to nearly 100% can be obtained from an aperture of only a few centimetres.

The next phase of the experiments was the testing of the theory.

A direct method of measurement was adopted, and an aperture similar to the hypothetical one described by Macfarlane was constructed. It consisted of an absorbing screen made of a sheet of a substance having a very low reflection coefficient, stuck to an aluminium plate. The non-reflecting side faced the launcher. The screen was suspended from a rope-and-pulley system so that the height above the guiding surface of its lower edge could be varied between zero and about 70 cm. The plane of the screen

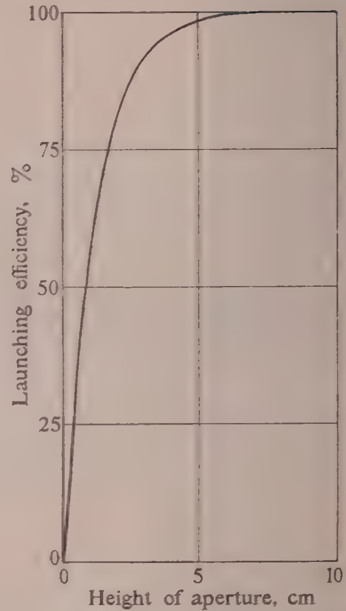


Fig. 4.—Launching efficiency of an aperture containing a field of curtailed exponential form.

was normal to the direction of propagation of the surface wave. It was 2 ft wide and 3 ft high, and so the effects of radiation leaking around it were neglected.

Imagine now a surface wave set up by the launcher impinging on this screen. Then all the energy will be absorbed except that portion which travels through the aperture, and the field in the aperture will thus be of the desired curtailed exponential form.

It has been assumed throughout the foregoing discussion that $H_x = \rho E_y/C_0$ in the aperture. This is true only if the screen is "black," i.e. that all radiation falling onto it is absorbed and that it causes no distortion of the electric or magnetic field in its vicinity. The use of any other type of screen would modify the fields and so invalidate the theory. The screen used in the experiments approximated closely to the theoretical ideal.

Even so, the former method is not applicable, for the essential step of matching the aperture to the surface, when the latter is terminated by the load, cannot be taken; any matching device would interfere with the exponential field distribution, and some of the energy would be scattered into a free-space wave and so lost to the system. Integration of the surface-wave field by a large and well-matched horn would be feasible only for large distances from the aperture where the surface wave predominates.

Owing to the difficulties in measuring the launching efficiency outlined above, what was finally measured to test the theory was the ratio of the fields in the aperture and at a distance beyond it. If E_r is the field strength at the surface at a distance r from the aperture and E_a that within the aperture, then from eqns. (16) and (21),

$$\frac{E_r}{E_a} = (1 - \varepsilon^{-2kXh})\varepsilon^{-jk\sqrt{(1+X^2)}z} \quad . \quad . \quad . \quad (33)$$

putting $S_0' = iX$ as before. The z-dependent factor, which concerns only the propagation of the wave, is neglected. The square of the above quantity is displayed at the output owing to the square law of the crystal, and so the quantity measured is

$$\frac{E_r^2}{E_a^2} = (1 - \varepsilon^{-2kXh})^2 \quad . \quad . \quad . \quad . \quad . \quad (34)$$

This is equal to the square of the launching efficiency.

* This is an approximate value, but the solution of eqn. (31) seems to be quite insensitive to changes in ϵ. A 4% change in ϵ produces a change of only $0\cdot0001$ in p_0. On the other hand, changes in a produce a much greater effect.

† The average thickness of the sheet was about $0\cdot006$ in over the nominal 1/16 in.

Measurements of E^2 were taken both within the aperture and at a distance of 86 cm beyond the aperture. This is the maximum distance allowed by the construction of the apparatus. In both positions there was a periodic variation of E as the probe was moved because of standing waves. The mean* value was taken in all cases.

It was found that effect of the power carried within the dielectric was small enough to be neglected.

There remains to be examined the effect upon the final results of the Fresnel integrals in eqn. (18). Because of the presence of these, the practical results may not be always in accordance with the simple formula of eqn. (34). Suppose that in eqn. (18) h is very large, then the second of the Fresnel integrals tends to zero and so does its coefficient ε^{-2kXh}, and the first Fresnel integral becomes simply $\sqrt{(2)}\varepsilon^{-j\pi/4}$. The field beyond the aperture is then

$$E_y = \varepsilon^{-jk\sqrt{(1+X^2)}z}\sqrt{(1+X^2)}\,\varepsilon^{-kXy} \quad . \quad . \quad (35)$$

which is the same as the field within the aperture. This means that the field distribution is unaffected when the wave travels through a very large aperture. Although this is intuitively obvious, the foregoing discussion justifies the correction of the experimental results described below. On the other hand, when h is very small both integrals reduce to $F[jX\sqrt{(\pi r/\lambda)}]$, and it has been assumed that r is large enough for this to be put equal to $\sqrt{(2)}\varepsilon^{-j\pi/4}$, which is the first term of the asymptotic expansion of eqn. (20). Putting $\lambda = \pi$ cm [this will simplify computation, and as $\lambda = 3{\cdot}15$ cm in actual fact only a very small error will result in $\sqrt{(\pi/\lambda)}$], $X = 1/5$, and $r = 85$ cm, it is found from eqn. (20) that $F \simeq 1{\cdot}58\varepsilon^{-j39^\circ}$, which differs from the first approximation by about 11% in modulus and 13% in argument. This discrepancy is not serious, for the efficiency itself will be small in this case.

For intermediate values of h the arguments of the Fresnel integrals are complex, and their asymptotic expansion becomes

$$F(u) = \sqrt{(2)}\varepsilon^{-j\pi/4} - \frac{j\varepsilon^{-ju^2}}{\sqrt{(2\pi)}u} \quad . \quad . \quad . \quad (36)$$

We must now use eqns. (16), (18) and (36) to get a more exact expression for E_r/E_a than the approximate eqn. (33). Neglecting the phase factor $\varepsilon^{-jk(1+X^2)z}$ and putting $y = 0$ (since all measurements are taken at the surface), we obtain

$$\frac{E_r}{E_a} = \frac{\varepsilon^{j\pi/4}}{\sqrt{(2)}}\left[F\left(\frac{-h}{\sqrt{r}} + jX\sqrt{r}\right) - \varepsilon^{-4hX}F\left(\frac{h}{\sqrt{r}} + jX\sqrt{r}\right)\right] \quad . \quad (37)$$

$\lambda = \pi$ cm as before, whence $k = 2$.

Using the asymptotic expansion [eqn. (36)] we have

$$\frac{E_r}{E_a} = \frac{\varepsilon^{j\pi/4}}{\sqrt{2}}\left[\sqrt{(2)}\varepsilon^{-j\pi/4} - \frac{j\varepsilon^{-j\left(\frac{h^2}{r} - X^2 r\right)}\varepsilon^{-2hX}}{\sqrt{(2\pi)}\left(-\frac{h}{\sqrt{r}} + jX\sqrt{r}\right)}\right.$$
$$\left. - \varepsilon^{-4hX}\sqrt{(2)}\varepsilon^{-j\pi/4} + \varepsilon^{-4hX}\frac{j\varepsilon^{-j\left(\frac{h^2}{r} - X^2 r\right)}\varepsilon^{2hX}}{\sqrt{(2\pi)}\left(\frac{h}{\sqrt{r}} + jX\sqrt{r}\right)}\right]$$
$$= 1 - \varepsilon^{-4hX} + \frac{\varepsilon^{j\pi/4}}{\sqrt{2}}\frac{\varepsilon^{j\pi/2}}{\sqrt{(2\pi)}}\frac{\varepsilon^{-j\left(\frac{h^2}{r} - X^2 r\right)}\varepsilon^{-2hX}}{(h^2/r + X^2 r)}\frac{2h}{\sqrt{r}} \quad . \quad (38)$$

* This is not strictly accurate, but for cases like this where the reflection coefficient is small the error is negligible.

Putting in the same values for the various constants as before and putting $h = 3$ cm, we have

$$\frac{E_r}{E_a} \simeq 1 - \varepsilon^{-4hX} + 0{\cdot}012(1 - j)$$

Thus the correction to E_r/E_a is of the order of 1·2% for $h = 3$ cm and the experimental error will be twice this since it is E_r^2/E_a^2 that is measured. This correction then will be neglected to a first approximation, but later it may be necessary to take account of it.

Figs. 5 and 6 show the results of the experiment. The ratio

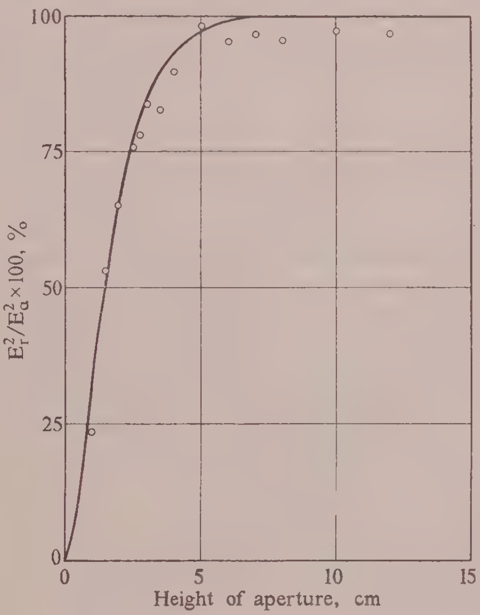


Fig. 5.—Ratio of squares of electric fields within the aperture and at a distance $r = 66$ cm beyond it.

$X = 0{\cdot}21$.
$\lambda = 3{\cdot}15$ cm.
$k = 2$.

——— Theoretical curve.
○ Experimental points (corrected).

of the squares of the fields is shown, for this makes comparison rather easier. The two graphs are for $r = 66$ cm and $r = 85$ cm respectively. The experimental points have all been corrected by multiplying by the same constant factor, so that when the screen is drawn right out of the way $E_r^2 E_a^2 = 100\%$. It can be seen that agreement between the experimental points and the curve $(1 - \varepsilon^{-2kXh})^2$ drawn in on the same scale is good for small values of h, but that the departures from the theory become significant for $h = 3$ cm. Now the errors introduced by the power carried within the dielectric are greatest for these small values of h, but the correction due to the Fresnel integrals may persist for values of h larger than 3 cm. The neglect of the power within the dielectric is thus justified to some extent by the experimental results.

The next step was to compute the relative field strengths taking account of the Fresnel integrals. Fresnel integrals for the values of complex arguments needed for this calculation are not tabulated, but there does exist a set of altitude charts for complex arguments of the error function, to which the Fresnel integral is closely related, prepared by T. Laible.[33]

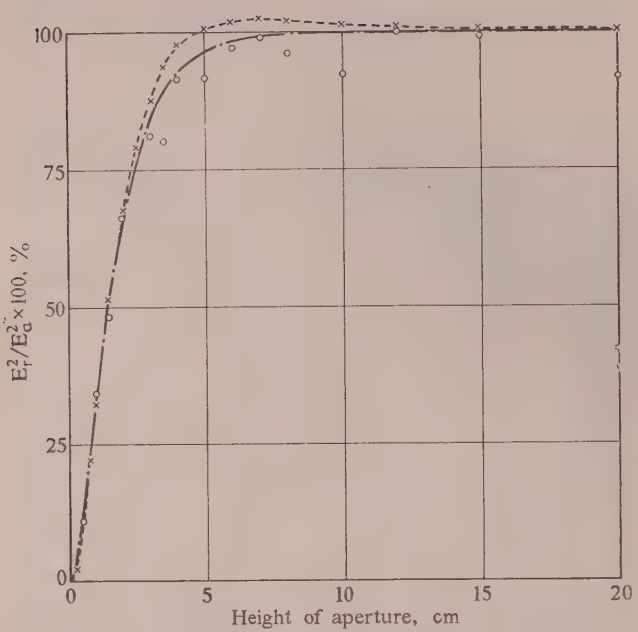


Fig. 6.—Ratio of the squares of electric fields within the aperture and at a distance $r = 85$ cm beyond it.

$X = 0{\cdot}21.$
$\lambda = 3{\cdot}15$ cm.
$k = 2.$

— · — Theoretical curve (to first approximation).
– – – – Theoretical curve (taking account of Fresnel integrals).
⊙ Experimental points.

The Fresnel integral discussed above is

$$F(u) = \sqrt{\left(\frac{2}{\pi}\right)}\int_u^\infty \varepsilon^{-j\zeta^2}d\zeta$$

whereas Laible plots

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z \varepsilon^{-\tau^2}d\tau \quad . \quad . \quad . \quad . \quad (39)$$

By splitting the Fresnel integral into $\int_0^\infty \varepsilon^{-j\zeta^2}d\zeta$ and $-\int_0^u \varepsilon^{-j\zeta^2}d\zeta$ and making the substitution $\zeta = \varepsilon^{-j\pi/4}\tau$, it can be shown that

$$F(u) = \frac{\varepsilon^{-j\pi/4}}{\sqrt{2}}[1 - \operatorname{erf}(u\varepsilon^{j\pi/4})] \quad . \quad . \quad . \quad (40)$$

Then the expression for the relative field strengths becomes

$$\frac{E_r^2}{E_a^2} = \frac{1}{4}\Bigg[(1 - \varepsilon^{-k2Xh}) - \Bigg\{\operatorname{erf}\frac{1}{\sqrt{2}}\Bigg[-\left(\frac{h}{\sqrt{r}} + X\sqrt{r}\right) - j\left(\frac{h}{\sqrt{r}} - X\sqrt{r}\right)\Bigg] - \varepsilon^{-2kXh}\operatorname{erf}\frac{1}{\sqrt{2}}\Bigg[\left(\frac{h}{\sqrt{r}} - X\sqrt{r}\right) + j\left(\frac{h}{\sqrt{r}} + X\sqrt{r}\right)\Bigg]\Bigg\}\Bigg]^2 \quad . \quad (41)$$

putting $\lambda = \pi$ as before.

The value of the above expression was computed for the usual range of values of h (0 to 20cm) and the results are plotted in the uppermost curve of Fig. 6. This curve approximates closely to the simple curve $(1 - \varepsilon^{-2kXh})^2$ for large and small values of h, but at intermediate values there is some discrepancy. The taking into account of the Fresnel integrals does not explain the deviations from theory of the experimental results, for the curve defined by eqn. (41) is further from the experimental points than that defined by eqn. (34). The new curve gives $E_r^2/E_a^2 >$ 100% for several values of h—a result which is not alarming since E_r^2/E_a^2 is not in this case equal to the square of the launching efficiency. But this might account for some of the high values of E_r^2/E_a^2 encountered in the measurements. So the Fresnel integrals will not account for the deviations of the experimental points from the curve $E_r^2/E_a^2 = (1 - \varepsilon^{-2kXh})^2$.

A closer examination of the field produced by the aperture was made. The method adopted was to take a series of measurements of the variation of field strength with height above the surface using the automatic recorder. A constant aperture with $h = 2{\cdot}5$ cm was employed and measurements were taken for various values of r.

Measurements according to the following system were made on the exponential portions of the various curves. The values of y at which the square of the field strength had fallen to $1/\varepsilon$, $1/\varepsilon^2$, and $1/\varepsilon^3$ of an arbitrarily chosen value were noted. The reference point was taken close to the point of inflexion of the curve to avoid the disturbing effect of the peak. The distances between the successive points were measured and called Δy_1,

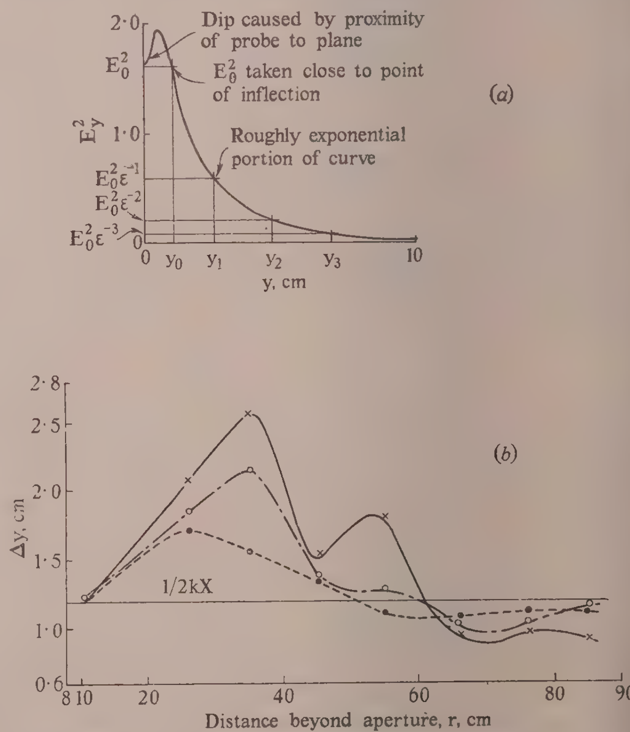


Fig. 7.—Curves of field strength and form of field.

(a) Variation of field strength with height above plane showing the method of analysing the graph.

$h = 2{\cdot}5$ cm.
$r = 26$ cm.
$\Delta y_n = y_n - y_{n-1}.$

(b) Variation of the form of the field with distance beyond the aperture.

● = Δy_1.
○ = Δy_2.
× = Δy_3.

Δy_2 and Δy_3. Difficulty of measurement precluded the taking of further points. The method is illustrated in Fig. 7(*a*).

Now, if the field strength conforms to the theory,

$$E_y \propto \varepsilon^{-kXy}$$

and the values of Δy will all be equal to $\frac{1}{2}kX$ (the factor 2 is introduced by the square law of the detector). If $k = 2$, $X = 0{\cdot}21$, then $\frac{1}{2}kX = 1{\cdot}2$cm. This distance will be referred to as the *space-constant* of the surface-wave mode by analogy with the time-constant of *RC* circuits.

In only one case do the three measured values of Δy equal $1{\cdot}2$cm; this is for $r = 11$cm, but for larger distances the field spreads out and the space-constant varies according to the distance from the surface. A graph showing the values of Δy for the values of r for which measurements were taken appears in Fig. 7(*b*), and tentative curves joining the various values of Δy_1, Δy_2 and Δy_3 are drawn. When these three curves are most widely separated, the field departs most widely from exponential form. For $r = 35$cm the field is very much spread out, but further from the aperture it becomes closer to the theoretical form, and for the greatest values of r it seems to approach the theoretical value of the space-constant asymptotically. It is significant that in all cases* it is Δy_1 that is closest to $1{\cdot}2$cm. This is the value taken nearest the surface, and it shows once more how the surface wave becomes dominant over the diffraction field near to the surface.

Similar measurements taken for the field from the surface-wave launcher show the same tendency; it is always closest to the surface that the surface wave is dominant. This is, of course, in line with the measurements of X which are in close accord with theory. The field from the variable aperture begins to settle down to the theoretical form when $r = 80$cm, but the diffracting screen is only 52cm from the end of the launching flare so that the field set up in the plane of the aperture will depart from the true exponential for quite small distances above the surface. For small values of h the aperture would contain a field of nearly pure exponential form, but as h increases the upper portions of the aperture will contain a field seriously modified by the diffraction pattern of the launcher. It has been assumed in all the foregoing theoretical discussion that the aperture contains a field of pure exponential form, but in practice this will be more and more seriously modified as h increases. It is to be expected that the launching efficiency will most nearly approach the theoretical value for small values of h. This would account for the discrepancies in the measured values of E_r^2/E_a^2 on the assumption that the distorted aperture field results in an increase in the diffraction field of the aperture at the expense of the surface-wave field. The strength of the surface wave at $y = 0$ would then be reduced, and the measured values of E_r^2/E_a^2 might resemble those shown in Fig. 6.

(7) CONCLUSIONS

The main object of the work described in the previous Sections was to confirm the reported properties of the surface wave and to investigate the setting-up of a plane surface wave. It was found to be relatively easy to excite a surface wave; almost any launching device would do that, and the surface wave when set up behaved much as was predicted in such matters as field distribution and phase velocity. The efficient excitation of the surface wave was more difficult, but the 50% efficiency of the double-cheese launcher, which was designed on purely empirical considerations, provided a surface-wave field of sufficient purity to enable a more systematic investigation to be undertaken. It was found that, in theory, it is possible to get launching efficiencies of nearly 100% from apertures no more than 2λ in extent, but that, in practice, such efficiencies were not attained. Thus it was presumed, after examination of various possible sources of error, that it is necessary to have a very good approximation to the exponential field distribution within the launching aperture. The design of such a launcher is evidently the next part of this problem to be attacked. For small apertures, which can have a nearly exponential field distribution, the theory was followed quite closely by the experimental results.

* There is one exception: for $r = 87$cm, Δy_2 is greatest, but this discrepancy is within the limits of experimental error.

(8) ACKNOWLEDGMENTS

Acknowledgment is made to the Chief Scientist, Ministry of Supply, and to the Controller of H.M. Stationery Office for permission to publish the paper.

The author wishes to express his thanks to three members of the staff of the Radar Research Establishment for their help and advice during the investigations: Mr. S. S. D. Jones under whose guidance the project was carried out; Dr. T. B. A. Senior for invaluable advice on diffraction problems and, in particular, the properties of Fresnel integrals; and Mr. J. D. Bradshaw, who constructed nearly all the special apparatus involved in the measurements.

(9) REFERENCES

(1) SOMMERFELD, A.: "Über die Fortplanzung electrodynamischer Wellen längs eines Drahtes," *Annalen der Physik und Chemie*, 1899, **67**· p. 233.
(2) ZENNECK, J.: "Über die Fortpflanzung ebener elektromagnetischer Wellen längs einer ebener Leiterfläche und ihre Beziehung zur drahtlosen Telegraphie," *Annalen der Physik*, 1907, **23**, p. 846.
(3) SOMMERFELD, A.: "Über die Ausbreitung der Wellen in der drahtlosen Telegraphie," *ibid.*, 1909, **28**, p. 665.
(4) SOMMERFELD, A.: "Über die Ausbreitung der Wellen in der drahtlosen Telegraphie," *ibid.*, 1926, **81**, p. 1135.
(5) WEYL, H.: "Ausbreitung Elektromagnetischer Wellen uber einen ebenen Leiter," *ibid.*, 1919, **60**, p. 481.
(6) ROLF, B.: "Graphs to Professor Sommerfeld's Attenuation Formula for Radio Waves," *Proceedings of the Institute of Radio Engineers*, 1930, **18**, p. 391.
(7) WISE, W. H.: "The Physical Reality of Zenneck's Surface-Wave," *Bell System Technical Journal*, 1937, **16**, p. 35.
(8) RICE, S. O.: "Series for the Wave-Function of a Radiating Dipole at the Earth's Surface," *ibid.*, 1937, **16**, p. 101.
(9) NORTON, K. A.: "The Physical Reality of Space and Surface Waves in the Radiation Field of Radio Antennas," *Proceedings of the Institute of Radio Engineers*, 1937, **25**, p. 1192.
(10) BURROWS, C. R.: "The Surface-Wave in Radio Propagation over Plane Earth," *ibid.*, 1937, **25**, p. 219.
(11) EPSTEIN, P. S.: "Radio-Wave Propagation and Electromagnetic Surface-Waves," *Proceedings of the National Academy of Sciences, Washington*, 1947, **33**, p. 195.
(12) BOUWKAMP, C. J.: "On Sommerfeld's Surface Wave," *Physical Review*, 1950, **80**, p. 294.
(13) KAHAN, T., and ECKART, G.: "L'onde de surface de Sommerfeld," *Comptes Rendus de l'Académie des Sciences (Paris)*, 1948, **226**, p. 1513.
(14) KAHAN, T., and ECKART, G.: "Sur le choix des chemins d'intégration dans le probléme relatif au rayonnement," *ibid.*, 1948, **227**, p. 969.
(15) KAHAN, T., and ECKART, G.: "La propagation des ondes électromagnétiques au-dessus du pol," *Le Journal de Physique et le Radium*, 1949, **10**, p. 165.
(16) KAHAN, T., and ECKART, G.: "On the Electromagnetic Surface Wave of Sommerfeld," *Physical Review*, 1949, **76**, p. 406.
(17) KAHAN, T., and ECKART, G.: "On the Existence of a Surface Wave in Dipole Radiation over a Plane Earth," *Proceedings of the Institute of Radio Engineers*. 1950, **38**, p. 807.

(18) KAHAN, T., and ECKART, G.: "Die Nichtexistenz der Oberflächenwelle in der Dipolstrahlung über einer ebener Erde," *Archiv der Elektrischen Ubertragung*, 1951, **5**, p. 25.
(19) KAHAN, T., and ECKART, G.: "Zur Frage der Oberflächenwelle in der Dipolstrahlung über einer ebenen Erde," with a reply by H. OTT, "Oberflächenwelle und kein Ende," *ibid.*, 1951, **5**, p. 347.
(20) OTT, H.: "Gibt es im Feld eines Senders eine Zenneckwelle?" *ibid.*, 1951, **5**, p. 15.
(21) GOUBAU, G.: "Surface-Waves and their Application to Transmission Lines," *Journal of Applied Physics*, 1950, **21**, p. 1119.
(22) ROTMAN, W.: "A Study of Single-Surface Corrugated Guides," *Proceedings of the Institute of Radio Engineers*, 1951, **39**, p. 952.
(23) REYNOLDS, D. K., and LUCKE, W. S.: "Corrugated End-Fire Antennas," *Proceedings of the National Electronics Conference*, 1950, **6**, p. 16.
(24) ERLICH, J. J., and NEWKIRK, L.: "Corrugated Surface Antennas," Hughes Aircraft Company Technical Memorandum No. 297, 1953. C.R.B. reference No. 53/1993.
(25) BARLOW, H. M., and CULLEN, A. L.: "Surface-Waves," *Proceedings I.E.E.*, Paper No. 1482 R, April, 1953 (**100**, Part III, p. 329).
(26) STRATTON, J. A.: "Electromagnetic Theory" (McGraw-Hill Book Company Inc., 1941), p. 516.
(27) ATTWOOD, S. S.: "Surface-Wave Propagation over a Coated Plane Conductor," *Journal of Applied Physics*, 1951, **22**, p. 504.
(28) SHERSBIE-HARVEY, R. B. R.: Some Preliminary Notes on Dielectric Guiding Surfaces for Electromagnetic Waves," T.R.E. Memorandum No. G10/157/RBRSH, 1945.
(29) MACFARLANE, G. G.: "Notes on Launching of a Zenneck Wave over the Surface of the Earth," T.R.E. Memorandum No. 472, 1951.
(30) CULLEN, A. L.: "The Excitation of Plane Surface Waves," *Proceedings I.E.E.*, Monograph No. 93 R, February, 1954 (**101**, Part IV, p. 225).
(31) KIELY, D. G., COLLINS, A. E., and EVANS, G. S.: "Cheese Aerials for Marine Navigational Radar," *ibid.*, 1951, **98**, Part III, p. 37.
(32) BOOKER, H. G., and CLEMMOW, P. C.: "A Relation between the Sommerfeld Theory of Radio Propagation over a Flat Earth and the Theory of Diffraction at a Straight Edge," *ibid.*, 1950, **97**, Part III, p. 18.
(33) LAIBLE, T.: "Höhenkarte des Fehler-integrals," *Zeitschrift für angewandte Mathematik und Physik*, 1951, **2**, p. 485.

621.385.16.029.6 : 621.317.6 : 621.317.361

Paper No. 1792 R
March 1955

COLD MEASUREMENTS OF 8MM MAGNETRON FREQUENCY AND PULLING FIGURE

By A. E. BARRINGTON, Ph.D., B.Sc., Associate Member.

(*The paper was first received 4th August, and in revised form 29th October*, 1954.)

SUMMARY

After a brief discussion of existing methods of measurement and their limitations, a simple arrangement is described which makes it possible to obtain the required parameters quickly and accurately.

(1) INTRODUCTION

In experimental as well as quantity production of magnetrons it is desirable and frequently essential to determine performance characteristics of valves so far as possible by means of cold measurements. These are carried out prior to final assembly, and greatly simplify modification of design and elimination of faulty components.

Since the multi-cavity magnetron is essentially a fixed-frequency device, the frequency has to be controlled within specified limits. Tuning methods have been developed, but these introduce constructional difficulties which are generally complex and become prohibitive at millimetre wavelengths. The frequency is affected by the characteristics of the load connected to the magnetron, and in general it is not possible to maintain these at a constant value, since the exact electrical length of waveguide attached to the magnetron is not under control. It is therefore convenient to define pulling figure as the sum of the maximum frequency excursions on either side of the nominal frequency due to changes of position of a reflecting discontinuity giving a 1·5 : 1 standing-wave ratio.

The value of the pulling figure is determined by the coupling of the magnetron into its waveguide, and it has been found to be a very difficult quantity to control—particularly in short-wavelength magnetrons. The coupling coefficient cannot be altered once the valve is completed.

In order to obtain cold measurements of frequency and pulling figure it is customary to measure the reflection of the cold block on the end of a length of waveguide with a variable-frequency local oscillator. The resonant frequency is indicated by minimum reflection. The sharpness of response in detecting resonance will depend on the loaded Q-factor of the block. Measurements of the Q-factor are important because—as well as determining the overall efficiency of the magnetron—the loaded Q-factor determines the magnitude of the effect on the oscillating frequency of changes in the impedance of the external load, and hence the pulling figure.

The most commonly employed methods to obtain resonant frequency and pulling figure are:

(*a*) Determination of Q-factors based on measurements of standing-wave ratio.[1,2] This involves observation of the variation of position of the minimum of the standing-wave pattern as a function of frequency. It requires a large number of experimental data and becomes lengthy and tedious when carried out with a manually controlled variable-frequency local oscillator and standing-wave indicator.

(*b*) Direct measurement of resonant frequency based on maximum power absorption at resonance.[3] This is a more rapid method which makes use of a frequency-modulated local oscillator coupled to a cathode-ray oscillograph time-base in such a way that the X-deflection of the spot varies with the generated frequency, giving a horizontal wavelength scale.

At resonance, the power reflected by the magnetron under test is a minimum. This can be determined by observing a dip in the amplitude of the reflected power as detected by means of a crystal in one arm of a directional coupler and displayed on the Y-axis of the cathode-ray-tube screen. The magnetron frequency can thus be measured directly with an absorption wavemeter. If the magnetron is set up with a 1·5 : 1 standing-wave-ratio movable puller, the whole constitutes a resonant system, and the pulling figure is merely the maximum change in resonant frequency for a change in puller position. This method can therefore be applied to the combined system of magnetron and puller to find the pulling figure. However, since the Q-factors of 8-mm magnetrons are comparatively small, the dip displayed on the cathode-ray-tube screen is not sufficiently sharp to give an accurate indication of the resonance position, and measurements are made more difficult owing to variations of local-oscillator output over the scanned frequency band. Whilst straightforward in theory, therefore, this method is not sufficiently accurate.

The simple arrangement described below makes it possible to overcome these limitations, since it is based on observations of the rate of change of standing-wave ratio, which, even with low Q-factors, can be measured quite easily at resonance. Furthermore, since standing-wave measurements are not affected by variations of the local oscillator output, this source of inaccuracy is eliminated.

(2) METHOD

The output from a square-wave amplitude-modulated continuous-wave oscillator is applied to the magnetron under test. Signals proportional to outgoing and reflected power are obtained by means of directional couplers and, after rectification and amplification, are applied to the X- and Y-plates of a cathode-ray tube. If signal A, proportional to the outgoing power, is applied to the X-plate, and signal B, proportional to the reflected power, to the Y-plate, a straight line trace is obtained which is inclined to the X-axis at an angle ϕ, where $\phi = \arctan B/A$. This angle is independent of the amplitude of the local-oscillator output and depends only on the impedance presented by the magnetron. At resonance ϕ is a minimum. By manual adjustment of the local-oscillator frequency, the minimum position can easily be observed on the cathode-ray-tube screen, since the rate of change of ϕ is a maximum at resonance;

i.e. if $$\tan \phi = \frac{B}{A} = x$$

then $$\frac{d\phi}{dx} = \frac{1}{\sec^2\phi} = \cos^2\phi$$

For small values of ϕ, $\cos \phi$ and hence $d\phi/dx$ is a maximum.

The resonant frequency can then be measured with an absorp-

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Dr. Barrington is at the Services Electronics Research Laboratory Extension (Admiralty).

*

tion wavemeter. Similarly, as indicated previously, by setting up the magnetron with a movable puller, the change of resonant frequency with change of position of the puller probe can be observed. The pulling figure is then obtained as the sum of the maximum frequency deviations from the nominal value of frequency as the puller probe is moved along the guide.

In the equipment used at present, as is shown in Fig. 1, a

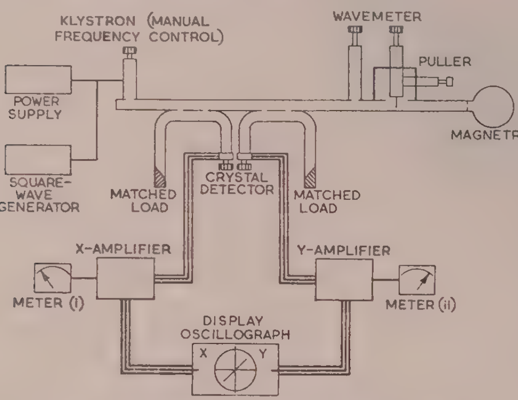


Fig. 1.—Eight-millimetre resonance display.

small phase difference between the X and Y amplifiers results in an elliptic rather than a straight-line display on the cathode-ray-tube screen. However, it is not difficult to observe resonance, which is indicated by an ellipse of minimum minor axis. Furthermore, if the power supplied by the klystron as read by meter (i) is kept reasonably constant, the reading on meter (ii), proportional to reflected power, indicates a sharply defined minimum at the resonant frequency.

(3) ACCURACY OF RESULTS

Cold-frequency determinations have been made to the same order of accuracy as those measured with a standing-wave indicator. Values of cold frequency were higher than the corresponding oscillating frequency by an almost constant difference of 120 Mc/s for a large number of valves. In the type of magnetron investigated, pulling figures of the order of 50 Mc/s are usually required. The measurement of the pulling figure of oscillating magnetrons depends critically on the standing-wave ratio of the test bench following the puller, and under optimum conditions it is accurate to within ± 5 Mc/s. Cold pulling-figure measurements on a number of magnetrons have given values which agreed to within this limit with the pulling figure of oscillating valves, a result which, in view of the change of nominal frequency in the two cases, may be considered somewhat surprising.

(4) ACKNOWLEDGMENTS

The paper is published with the approval of the Admiralty.

(5) REFERENCES

(1) Collins, G. B.: "Microwave Magnetrons" (McGraw-Hill Book Co., 1948), p. 713.
(2) Latham, R.: "Cold Impedance Measurements on Magnetrons" (B.T.H. Research Laboratory: CVD Report L320-S, 1945).
(3) Latham, R., King, A. H., and Rushforth, L.: "The Magnetron" (Chapman and Hall, 1952), p. 38.

DISCUSSION ON "PRINTED AND POTTED ELECTRONIC CIRCUITS"*

SOUTH MIDLAND RADIO GROUP, AT BIRMINGHAM, 23RD NOVEMBER, 1953

Mr. A. Bikker: During the last war the Germans found themselves very short of copper and began to use aluminium conductors. During war-time the question of how much can be got from every ton of raw material is vital, and it occurs to me that printed circuits would show a better material yield than conventional methods. Can the authors confirm this?

Mr. L. G. Ward: Dip soldering has been common practice for years on small motors for the soldering of squirrel cages and commutator connections. A rather limited experience of it has shown me that, although it is 100% reliable under ideal conditions, it can produce dry joints amongst the good ones if the conditions are not ideal.

Since the embedded resistors are level with the surface of the block, presumably there is only one surface from which the major part of the heat can be dissipated. It is possible that over this there may be a stationary film of air, because it is flush with the face of the block; if this is so, it could further reduce the cooling. Could the authors give an idea of the size of these resistors for a given heat dissipation?

Mr. A. W. Binns: As one who is mainly concerned with the domestic-radio field, I do not agree with the statement in Section 2 that the printing of grade II resistors, low-value capacitors, certain inductors and potentiometers and all the wiring "will be a major contribution, since approximately 50–70% of the total components in the average radio or radar equipment is composed of these items." Is this statement based on the monetary value or the number of components? I would agree with reference to a radio set if it is based on the number of components, but it is not true of the value, which lies almost entirely in a few large items such as the cabinet, the loudspeaker, ganged variable capacitors, mains transformers, etc., and until it became feasible to print these, the printed-circuit technique would not represent a great saving in cost. The position would probably be eased if we persuade the B.B.C. to use higher frequencies.

Has any attempt been made to produce or print variable capacitors or trimmers, or capacitors with values up to 450 $\mu\mu$F?

Brigadier F. Jones: Should we be right in concluding that the application of printed and potted circuits is more in the field of high frequency than for medium or low frequency, because of the smallness of the components?

Mr. M. P. Johnson: Can the authors say what an American factory turning out this equipment looks like? Do they have production lines and do they look like an English factory? What would be the American percentage of rejects?

Mr. T. D. G. Wintle: Can the authors give some idea of the durability of the printed film, particularly regarding its resistance to abrasion? It would appear that this means of connection could be used in electrical assessories and similar light-current applications. It would be undesirable to use a chemical process, but some form of offset printing would be economical. Any form of varnishing for protection would make the idea less attractive from the cost aspect.

* Dummer, G. W. A., and Johnston, D. L.: Paper No. 1407 R, November, 1952 (see 100, Part III, p. 177).

What maximum currents have been carried by printed circuits, and have clamped connections, with no soldering, been successful?

Mr. G. L. G. Jeans: In all the circuits shown in the paper there are no cross-overs. There may be circuits such as bridge circuits, where it is necessary to have cross-overs. How are such cross-overs insulated?

Mr. J. J. E. Aspin also contributed to the discussion.

Messrs. G. W. A. Dummer and **D. J. Johnston** (*in reply*): For convenience we have grouped our reply under the speakers' names.

Mr. A. Bikker.—It is possible that printed circuits can show a saving of material over conventional methods. The current-carrying capacity of the foil is high and the whole of one surface is exposed for heat dissipation. A normal connecting wire has a much greater volume than the equivalent printed-circuit wire. Although excess copper is bonded to the material, in the acid-etch system the copper can be recovered during the etching process.

Mr. L. G. Ward.—Experience with dip soldering has shown that if the conditions are ideal 100% reliability is obtained and it is stressed that mechanization of the process is essential. If this is done, then close control can be maintained and no failures should be experienced.

Although the embedded resistors are level with the surface of the block, so that one surface only is available for dissipating heat, the surface/thickness ratio is extremely high and, in general, carbon-mix resistors can be rated at up to $1\frac{1}{2}$ watts/in^2, this being limited by the nature of the material on which the resistor is printed.

Mr. A. W. Binns.—The figure of 50–70% of the total components is based not on the monetary value, but on the number of components. The greatest saving in cost using normal components is probably in the dip-soldering technique, which enables much hand-labour to be replaced. Some components, such as transformer windings, simple switches, etc., have been produced experimentally by printed means, and it is possible to produce small trimmer capacitors by rotational adjustment or by rotation of a small plate above the deposited metal, but no successful variable capacitor or high-value capacitors have yet been produced by printed techniques.

Brigadier F. Jones.—The limitation referred to in printed circuits would probably be in coils, where the restriction on the number of turns which can be printed limits the frequency range, although it should be pointed out that losses in the material on which the coil is printed will limit the upper frequency range.

In potted circuits the losses in the resins will most certainly limit its usefulness at high frequencies.

Mr. M. P. Johnson.—American factories producing printed circuits by standard techniques are similar to those in the United Kingdom, but most of the effort in America is being devoted to automatic assembly of electronic units; in this case the factories are laid out on completely different lines, and normal production lines of operatives are being replaced by static machines with a few attendants.

It is very difficult to estimate the American percentage of rejects. They are probably higher at the moment than with normal established techniques, but no doubt will be reduced as further experience is gained in the operation of these machines.

Mr. T. D. G. Wintle.—The durability of printed wiring on the etched-foil system is good, comparable with that of copper itself, and it is in use in some electrical accessories.

The maximum currents carried by printed circuits are very high indeed, and currents of 6 amp can be carried on 1·5 mil foil 1/32 in wide, with a temperature rise of 25° C above ambient, on a plastic base. On ceramic bases the current-carrying figures would be many times higher. So far as is known, no clamped connections have been tried. Problems here would be similar to those experienced between metals such as copper and the clamp material.

Mr. G. L. C. Jeans.—In most circuits cross-overs can be avoided by careful design, but where they are essential, eyelets or studs can be used with part of the wiring on the reverse side of the plate.

DISCUSSION ON "ELECTRONIC TELEPHONE EXCHANGES"*

MERSEY AND NORTH WALES CENTRE, AT LIVERPOOL, 4TH JANUARY, 1954

Mr. T. A. P. Colledge: At the present time we have no experience in the maintenance of electronic exchanges, but this will bear some similarity to maintenance in repeater stations in carrier and coaxial-cable systems. An analysis of faults found on overhauls of an automatic exchange and on these repeater stations shows that 60–80% of the faults are due to dry joints. This type of fault will occur in both electronic and electro-mechanical exchanges, so we can regard it as a common factor and ignore it when assessing the advantage of the one exchange over the other. Can the author confirm that this is a reasonable assumption? The remainder of the faults are due to electrical and mechanical troubles. In purely electronic equipment functional testing will be a matter of checking meter readings, which may be quickly done. Fault location is generally the longest operation and requires top-grade staff. Even so, a mechanical fault often passes a number of functional routine tests but still occurs intermittently in service. To overhaul a 2-motion selector thoroughly takes a skilled technician from 1 to $1\frac{1}{2}$ hours, whereas an electrical component can be found and replaced within a few minutes.

In complex equipment, poor design can make the changing of a component a major operation comparable with component changes in mechanical switches. Jacking-in blocks of equipment is desirable, but if taken to excess the jack points tend to become the major source of faults. Valve-holders have given considerable trouble in the past. Have we developed a valve-holder of a reliable type, or do we have to solder in our valves?

Again, when fast-operating electronic equipment is used, the designer is tempted to utilize common equipment to such a degree that there is no latitude for the maintenance staff to take it out of service for overhaul during the day. This concentrated use will undoubtedly save initial costs, but it may well eventually cause a loss of traffic and the good will of the public, apart from the extra maintenance costs involved. It is hoped designers will bear this in mind when deciding on the type and quantity of electronic equipment they put into service.

I have no figures for the clear division of faults found on an electro-mechanical equipment, but as a result of the study of fault reports I am convinced that 75% are due to mechanical failures. Could the author give figures of the comparative nature of faults found on that part of the Richmond exchange which is working on electronic equipment and the faults found

* FLOWERS, T. H.: Paper No. 1266, February, 1952 (see 99, Part I, p. 181).

on the electro-mechanical part of the same exchange? I understand the number of faults on the electronic exchange has fallen by about 10% since the valves have been left in position and not removed for testing purposes. Some day perhaps we may be able to replace the telephone dial at the subscriber's end of the system. I imagine this will be our next problem.

Mr. J. A. Mason: Two years have passed since the paper was first published, and of recent developments the author has referred to transistors, but what of magnetic-drum storage? Have these two items, and possibly others, affected the author's concept of the most promising switching system? The bulk of the post-war products of the manufacturers of electro-mechanical systems have been supplied to oversea customers against world-wide competition, and it was rather disturbing to note the statement in Section 8 that "the consequential possibilities in transmission network design and construction are such as to justify electronic-exchange development, even if the other signs were not so propitious." Improvements in transmission networks can have only long-term value, and oversea customers are vitally interested in first costs of exchange equipment. Apprehensions as to first costs are not lessened by the amount of equipment required (according to Fig. 7) for the subscriber's line circuit—that most vital of all parts of the switching equipment so far as economics are concerned.

The future of electronic switching systems must lie in the availability of reliable components and in economics in terms both of first costs and annual charges. The cost of maintenance is obviously important, but in oversea competitive markets the first cost determines whether or not one receives the contract for the exchange equipment.

Since electronic switching stems from techniques developed during the war for electronic computers, etc., international patent problems are likely to provide a serious obstacle to the development and employment of an ideal electronic switching system. Has the author any comments on this aspect?

Mr. L. J. Murray: There are two schools of thought in the design of electronic switching systems, one of which may be referred to as the "multipled element" switch school and the other as the multiplex school. From the paper one may get the impression that to form an electronic system of the multipled-element type one has to replace every switching contact in the existing Strowger system by a cold-cathode tube or other electronic element. That is not altogether the case. One can take a simple 100-line electro-mechanical exchange having, say, ten line-finders and ten 100-point final selectors, i.e. 20 switches each with 100 sets of switching-bank contacts, a total of 2 000 in all. Its electronic counterpart of the multipled-element type would also have ten 100-point line-finders and ten 100-point final selectors, but instead of two separate sets of bank multiples it is necessary to have only one, because the electronic switching system can, in effect, have the line-finder and final-selector wipers working on the same bank. Therefore there are only ten sets of contacts, making 1 000 in all, and so there is only half the total number of switching contacts as in the electro-mechanical system. This would not apply in larger exchanges, but for such exchanges there are switching systems which are becoming familiar through the present-day crossbar systems known as co-ordinate or link systems, and they can reduce the number of contacts to one-third or one-quarter of the number required in the Strowger switching system. Electronic elements such as cold-cathode tubes can be used to simulate these crossbar switches.

Referring to another trunking point, I have always thought that the old manual switchboard system gave an efficient trunking system, because a set of three operators working on one appearance of a 10 000-line multiple formed a switching unit with, say, 500 inlets, 10 000 outlets and perhaps 50 cord circuits as intervening links. That, in fact, was a form of multiplex switch, and electronic multiplex circuits are trying to do the same, but it would appear that for time-division multiplex a capacity of 100 outlets from the switch seems to be the practical limit. Does the author agree with this, because it makes it necessary to build up the exchange as in the present system by having several stages of switches? It appears that modulation and demodulation will be required at each stage, and this tends to offset the other advantages of the multiplex switch.

One point which the author has not mentioned is that the speed of electronic switching enables the best advantage to be taken of what is known as the second-trial feature, i.e. if the control cannot complete the call within a given time it makes a further effort. It has proved extremely useful in relay crossbar switching in the United States, and it is recognized that it reduces the number of subscribers' complaints.

Mr. J. B. F. Williams: One of the main points from the overall cost aspect is that for the first time in an automatic telephone exchange the "go" and "return" paths have been split. This means that we can have any reasonable amount of amplification in a subscriber's line, and therefore the very considerable cost of the external line plant may be reduced. At present we install progressively heavier cables of $6\frac{1}{2}$–10 lb gauge or larger from the exchange to the exchange boundary; with an electronic exchange we could, in fact, utilize a cable of $2\frac{1}{2}$ lb gauge throughout (except for the final terminations), and therefore in the British Post Office, at least, where we consider the overall economics of both the equipment and the line plant before any new switching is brought into being or old exchanges are modified, there must be a very considerable saving that could be made on line plant which would allow the electronic exchange equipment to be more expensive than the electro-mechanical equipment, while at the same time the electronic system as a whole could be more economic.

Mr. T. H. Flowers (*in reply*): I agree with Mr. Colledge that unintentional dry joints are a common factor at present, but one which I feel must be eliminated from all types of exchange. Intentional dry joints in the form of plugs and jacks should also be reduced to a minimum; certainly valve-holders should be replaced by soldered connections. It is not possible to quote reliable fault figures for the Richmond electronic directors, as these are experimental to gain experience in design and maintenance techniques, and as a result of which better equipment can now be made. It can be said, however, that the electronic directors as tested are at least as reliable as the electro-mechanical ones.

In reply to Mr. Mason, it is true that much has been disclosed and learnt since the paper was written. The general tendency has been to strengthen the belief that the fully electronic exchange will eventually compete with existing types of exchanges without changing the author's view that time-division multiplex is likely to be a very powerful tool in achieving this end. One can foresee that there may be trading and patent difficulties, but at this stage it is too early to try to assess them.

Mr. Murray's point about minimizing the number of multipled elements is no doubt true, but the number is still formidable. The maximum number of channels in a time-division multiplex is still about 100, but systems which were under developemnt when the paper was written make it possible to have much greater numbers of lines connected to one multiplex system, and thus to have connections through exchanges with only one or two modulation and demodulation stages.

I entirely agree with the merits of the second-trial feature, and also with Mr. Williams's points regarding exchange-line-plant construction and economics.

621.372.8 : 621.357.6/.8

Paper No. 1785 R
Mar. 1955

SURFACE ROUGHNESS AND ATTENUATION OF PRECISION-DRAWN, CHEMICALLY POLISHED, ELECTROPOLISHED, ELECTROPLATED AND ELECTROFORMED WAVEGUIDES

By J. ALLISON, B.Sc.(Eng.), Graduate, and F. A. BENSON, M.Eng., Ph.D., Associate Member.

(*The paper was first received* 20*th August, and in revised form* 15*th October*, 1954.)

SUMMARY

Detailed examinations of certain 3 cm waveguides have shown that the surface finish of precision-drawn tubing as manufactured at present is quite adequate for most applications. Such surfaces, however, may be improved, if desired, by careful chemical or electrolytic polishing or electroplating in bright baths under closely controlled conditions.

Some information on the surface finish of copper guides electroformed on various types of mandrel is also presented.

Formulae for calculating the attenuation of any H or E mode in a rectangular waveguide, so as to take account of surface roughness, have been developed from the original expressions derived by Kuhn.

A method is given for determining the actual value of attenuation in a waveguide sample without having to make careful measurements with elaborate equipment on long specimens.

A description is included of a new and simple technique, involving electropolishing, for examining the internal surface finish of waveguides; the method cannot, however, be used successfully on silver-plated sections.

LIST OF PRINCIPAL SYMBOLS

a = Short internal dimension of waveguide.
b = Long internal dimension of waveguide.
c = Velocity of light.
H_w = Magnetic field of a wave sliding on the inner surface of the waveguide wall.
K_{T1}, K_{T2}, K_p = Surface-roughness factors in attenuation formulae.
W' = Characteristic density of energy.
σ = Conductivity of waveguide wall metal.
μ_1 = Permeability of waveguide wall metal.
λ_g = Guide wavelength.
λ_e = Wavelength in unbounded dielectric.
λ_{cr} = Critical guide wavelength.
ϵ = Permittivity of dielectric (in this case air).
μ = Permeability of dielectric (in this case air).
α = Attenuation produced by wall metal.
α_H = Attenuation produced by wall metal for an H wave, neglecting surface roughness.
α_E = Attenuation produced by wall metal for an E wave, neglecting surface roughness.
α'_H = Attenuation produced by wall metal for an H wave, taking account of surface roughness.
α'_E = Attenuation produced by wall metal for an E wave, taking account of surface roughness.
Z_g = Specific impedance of waveguide.
Z_a, Z_b, Z_{cr} = Transverse impedances of waveguide.

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
Dr. Benson and Mr. Allison are in the Department of Electrical Engineering, University of Sheffield.

(1) INTRODUCTION

Abnormally high attenuation in certain conductors at microwave frequencies has been reported by Vivian,[1] Morgan,[2] Maxwell[3] and others.[4–10] Two papers concerned with waveguide attenuation have also been published recently by one of the authors.[11,12] One[11] deals with the attenuation in drawn rectangular tubes, giving maximum values of roughness to be expected in these tubes, and concludes that the discrepancies between measured and calculated attenuation values are due solely to the roughness of the internal surfaces. The second paper[12] describes a study that has been made of the roughness of some electroplated waveguides. Although silver-plating of waveguide components seems to be standard practice, it has been shown that plating may not give the expected reduction in attenuation and is probably not worth while for many applications.

The present paper describes detailed examinations of the roughness of various waveguide surfaces, the majority of which are much smoother than those previously tested.[11,12] The results of a few attenuation measurements on some precision-drawn waveguides have already been given, and the surface roughness of this type of tube has now been determined. The effects of chemical and electrolytic polishing on the nature of the surface finish have been observed and some further work on electroplating from bright baths has been undertaken. Details are also included of the surface roughness of copper guides electroformed on various types of mandrel.

It has been shown previously[11] that the attenuation produced in a rectangular waveguide carrying the normal H_{01} mode can be calculated from a modified Kuhn formula[13] by introducing three surface-roughness factors. In deriving this formula, however, it was assumed that for the top and bottom walls of the tube the transverse surface irregularities affect equally the losses due to transverse and longitudinal components of current. An improvement in the previous calculation procedure has therefore now been made without this assumption and has led to a more accurate expression. The calculations have also been extended to provide modified Kuhn formulae, taking account of surface roughness, for determining the attenuation of any H_{nm} or E_{nm} wave in a rectangular waveguide.

Results of attenuation measurements agree very well with the new modified theory. The more elementary theory previously submitted also gave a reasonable interpretation of the experimental figures because the values of the surface-roughness factors are not very different either from each other or from unity.

Morgan[2] has made a theoretical investigation of the power dissipated by eddy currents in a metallic surface at microwave frequencies in the presence of regular parallel grooves or scratches whose dimensions are comparable to the skin depth. The solution of this highly idealized problem shows that the power dissipation in grooves of various sizes and shapes transverse to the direction of induced-current flow is increased by about 60%

over its value for a smooth surface when the r.m.s. deviation of the grooved surface from an average plane is equal to the skin depth. The exact shape of the grooves, according to Morgan, is not critical. Loss caused by grooves parallel to the current flow is shown, in a particular case, to be about one third as great as the increase caused by transverse grooves of similar size. The results given herein indicate that the special kind of surface studied by Morgan is seldom, if ever, found in practice, a fact which has also been confirmed by earlier attenuation measurements.[11,12]

In the past it has been found difficult to determine the surface-roughness factors accurately, but a method has now been evolved for obtaining their precise values. In this way, the difference between actual and theoretical attenuation values can be predicted by calculation without the necessity of making careful measurements on long lengths of tube by means of special test equipment. In previous calculations the longitudinal surface-roughness factors have sometimes been taken as unity, but during the investigations here described they have always been measured with a view to greater accuracy in the calculations.

A technique involving electropolishing for examining the internal surface finish of waveguides is described. This method is a considerable improvement over the earlier ones and enables scratch-free surfaces to be obtained with very little effort. The results indicate that, for most applications, the surface finish of brass and copper precision-drawn tubes as at present manufactured is quite satisfactory, although chemical and electrolytic polishing or bright plating, under closely controlled conditions, will give a lower attenuation if this is desired.

(2) CALCULATION OF ATTENUATION IN WAVEGUIDES

A general formula for calculating the attenuation produced by losses in the wall metal of a waveguide has been derived by Kuhn[13] and can be applied to any wave mode in either rectangular or circular waveguides. The assumptions made by Kuhn in developing the formula have been previously stated[11] and it has been shown that, for the H_{01} mode, which is the one most commonly generated, the attenuation expression can be modified to account for surface roughness. This is done by introducing three additional factors, K_{T1}, K_{T2} and K_p. The factor K_{T1} is defined as the ratio of the length of the actual surface to that of the ideal surface for the long sides of the waveguide transverse to the axis, and K_{T2} is the corresponding factor for the short sides. The factor K_p accounts for the effects of the surface roughness in the longitudinal direction and is defined in a similar manner to K_{T1} and K_{T2}. Measurements show that for any given waveguide both the long and short sides have about the same values of K_p, so one longitudinal roughness factor only is introduced into the calculations.

The modified formula can be further improved as shown in the Appendix (Section 11), and for the particular H_{01} mode the new expression is

$$\alpha' = \left[\frac{c}{\sigma}\frac{\mu_1}{\mu}\left(\frac{\epsilon}{\mu}\right)^{1/2}\right]^{1/2}\frac{\lambda_g}{b(\lambda_e)^{3/2}}\left[\left(K_{T2}+\frac{b}{2a}K_{T1}\right)\frac{\lambda_e^2}{\lambda_{cr}^2}+K_p\frac{b}{2a}\left(1-\frac{\lambda_e^2}{\lambda_{cr}^2}\right)\right] \quad . \quad . \quad . \quad . \quad (1)$$

It is shown in the Appendix that the formulae derived by Kuhn for the attenuation of either an H_{nm} wave or an E_{nm} wave in a rectangular guide may also be modified to include the effects of surface roughness. The expression for the H_{nm} wave then becomes

$$\alpha'_H = \left[\frac{c}{\sigma}\frac{\mu_1}{\mu}\left(\frac{\epsilon}{\mu}\right)^{1/2}\right]^{1/2}\frac{\lambda_g K_r}{b(\lambda_e)^{3/2}}\left(\frac{\lambda_e^2}{\lambda_{cr}^2}+K_t\right) \quad . \quad . \quad (2)$$

where

$$K_r = \frac{\left(\frac{2a}{n}\right)^2(aK_{T2}+bK_{T1}-bK_p)+\left(\frac{2b}{m}\right)^2(aK_{T2}+bK_{T1}-aK_p)}{a\left[\left(\frac{2a}{n}\right)^2-\left(\frac{2b}{m}\right)^2\right]} \quad . \quad . \quad . \quad . \quad (3)$$

and

$$K_t = \frac{bK_p\left(\frac{2a}{n}\right)^2+aK_p\left(\frac{2b}{m}\right)^2}{\left(\frac{2a}{n}\right)^2(aK_{T2}+bK_{T1}-bK_p)+\left(\frac{2b}{m}\right)^2(aK_{T2}+bK_{T1}-aK_p)} \quad . \quad . \quad . \quad . \quad (4)$$

An alternative form of eqn. (2), in which the ratio λ_e/λ_{cr} is the only parameter involving wavelengths, is sometimes useful; it is

$$\alpha'_H = \left[\frac{c}{\sigma}\frac{\mu_1}{\mu}\left(\frac{\epsilon}{\mu}\right)^{1/2}\right]^{1/2}\frac{1}{b^{3/2}\left[\frac{\lambda_e}{\lambda_{cr}}\left(1-\frac{\lambda_e^2}{\lambda_{cr}^2}\right)\right]^{1/2}}\delta K_r\left(\frac{\lambda_e^2}{\lambda_{cr}^2}+K_t\right) \quad (5)$$

where
$$\delta = \left(\frac{m}{2}\right)^{1/2}\left[1+\left(\frac{b}{a}\right)^2\left(\frac{n}{m}\right)^2\right]^{1/4} \quad . \quad . \quad . \quad (6)$$

The expression for the E_{nm} wave becomes

$$\alpha'_E = \left[\frac{c}{\sigma}\frac{\mu_1}{\mu}\left(\frac{\epsilon}{\mu}\right)^{1/2}\right]^{1/2}\frac{1}{b}\frac{\lambda_g}{(\lambda_e)^{3/2}}K'_r K_p \quad . \quad . \quad (7)$$

where
$$K'_r = \frac{1+(b/a)^3(n/m)^2}{1-(b/a)^2(n/m)^2} \quad . \quad . \quad . \quad . \quad . \quad (8)$$

With the ratio λ_e/λ_{cr} as the only parameter involving wavelengths the alternative form of this expression is

$$\alpha'_E = \left[\frac{c}{\sigma}\frac{\mu_1}{\mu}\left(\frac{\epsilon}{\mu}\right)^{1/2}\right]^{1/2}\frac{1}{b^{3/2}\left[\frac{\lambda_e}{\lambda_{cr}}\left(1-\frac{\lambda_e^2}{\lambda_{cr}^2}\right)\right]^{1/2}}\delta K'_r K_p \quad . \quad (9)$$

Thus the ratio of the calculated attenuation which takes account of surface roughness to that which assumes a perfectly flat surface becomes, for the H_{nm} wave,

$$\frac{\alpha'_H}{\alpha_H} = \frac{K_r\left(\frac{\lambda_e^2}{\lambda_{cr}^2}+K_t\right)}{K'_r\left(\frac{\lambda_e^2}{\lambda_{cr}^2}+K'_t\right)} \quad . \quad . \quad . \quad . \quad (10)$$

where
$$K'_t = \frac{b}{a}\left[\frac{1+(b/a)(n/m)^2}{1+(b/a)^3(n/m)^2}\right] \quad . \quad . \quad . \quad . \quad (11)$$

The expression for E_{nm} waves corresponding to eqn. (10) is

$$\frac{\alpha'_E}{\alpha_E} = K_p \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

For the special case of the H_{01} mode

$$\frac{\alpha'_H}{\alpha_H} = \frac{\left(K_{T2}+\frac{b}{2a}K_{T1}\right)\frac{\lambda_e^2}{\lambda_{cr}^2}+\frac{bK_p}{2a}\left(1-\frac{\lambda_e^2}{\lambda_{cr}^2}\right)}{\frac{\lambda_e^2}{\lambda_{cr}^2}+\frac{b}{2a}} \quad . \quad . \quad (13)$$

Thus if the values of the surface-roughness factors are determined for a specific guide, its actual attenuation for any mode can be accurately estimated from the above equations. The

values of K_{T1}, K_{T2} and K_p alone are useful for comparing the practical performance of various surfaces without having to use the equations.

(3) ELECTROPOLISHED WAVEGUIDES

For some metals, electropolishing provides a remarkably good surface finish, at least equal to, and in some cases better than, that given by the usual mechanical processes. The principle underlying electropolishing is that of selective dissolution under the influence of current, whereby the surface is made progressively smoother and more brilliant. The actual mechanism is still obscure, but the presence of a viscous film of the products of interaction between the metal and the electrolyte at the metal/liquid surface is known to be responsible for the uneven dissolution of the anode surface.[14] The film/electrolyte junction does not follow the metal surface irregularities so that the film thickness is not constant. Jacquet[15] explains that the film influences the dissolution of the anode because the resistance of the film is lower at the crests than at the hollows. This results in a varying current-density and preferential dissolution which leads to a general smoothing of the surface. Other theories[16,17] indicate that the polishing mechanism is closely associated with diffusion of the dissolution products through the concentrated layer.

Many metals can be polished in this manner[14] and there is no difficulty in treating copper, brass, silver or aluminium waveguides. Chemical and structural inhomogeneities constitute serious difficulties in electropolishing, but since waveguides are generally made of either pure metals or simple alloys such problems are not so important.[18] Further, it is the microprofile of a surface which is eliminated by electropolishing, the macroprofile, or general waviness, remaining. This is an advantage when polishing waveguides, since the microroughness affects the attenuation more than the general undulations[2] do.

During the investigations described a 2in length of 3cm precision-drawn copper guide was polished internally, using a smaller rectangular copper waveguide as an internal cathode and the specimen as anode. The polishing bath consisted of two parts of orthophosphoric acid and one part of water. The voltage, which was found to be very critical, was 1·8 volts and the polishing time was 30min. At voltages below the optimum value the specimen presented an etched appearance; higher voltage caused gas evolution and pitting.

A 2in length of 3cm precision-drawn brass waveguide was next electropolished internally, using a 0·75in × 0·25in solid copper cathode in the same bath as used for the copper sample. This cathode had a much better throwing power than the previous one. It was also found that the optimum value for the polishing voltage can be deduced from the voltage/current characteristic of the electrolyte. The best polish is obtained at a voltage just below that which causes gassing at the anode, as indicated by a sharp increase in current density for a small change in voltage. By this method the smoothest internal finish was found to be produced by a voltage of 2·45 volts and a polishing time of 50min.

(4) CHEMICALLY POLISHED WAVEGUIDES

Chemical polishing has been used in an attempt to improve the interior finish of both brass and copper drawn waveguides. The theory of the levelling produced by this method has been given by Pinner.[19] A thin film of oxide or basic salt of the metal is present on the surface during polishing. The surface is smoothed because slower diffusion of metal ions from microdepressions than from microelevations causes the film over the elevations to be dissolved at a greater speed by anions in the polishing solution. The metals which can be chemically polished are copper, brass,[19] bronze, nickel–silver, nickel, zinc, aluminium, iron, steel and most aluminium alloys.[20,21] Bright-dipping processes for copper, most of which are based on mixtures of nitric acid, sulphuric acid and water, have been used with some success, and a definite smoothing action has been achieved.[22] Chemical polishing solutions based on phosphoric-acid mixtures have a similar overall smoothing action for large surface irregularities, but they have a higher micropolishing action which produces a finish of considerable specular reflectivity. These solutions are suitable for brass, nickel–silver, nickel and Monel metal. To obtain the desired result the total thickness of metal removed must be between 0·001 and 0·0015in. Best results are obtained, as in electropolishing, with pure metals and single-phase alloys. As a general rule, pronounced differences may be found in the finishes obtained on alloys of the same composition, the degree of smoothness varying with the grain size. Cold-worked surfaces polish better than annealed sheet or strip. Provided that the correct operating conditions and composition of the solution are found, chemical polishing can be used to improve the internal surface finish of copper and brass waveguides.

Chemical polishing of some precision-drawn copper guide was carried out during the investigations here described in a commercial phosphoric acid solution, the exact contents of which could not be ascertained. Some precision-drawn brass guide was then polished by a process and with a polishing solution identical with those employed for the copper. In this way, comparison could be made between chemical polishing of surfaces of different composition and texture.

(5) BRIGHT-PLATED SURFACES

(5.1) Silver

The addition of carbon disulphide to silver-cyanide baths as a brightening agent was suggested as early as 1847. More modern addition agents contain sulphur in the form of turkey-red oil,[23] cyanates, etc. The addition of these agents in the correct proportions gives a considerably brighter silver surface with a certain degree of levelling. Previous results[12] on a waveguide plated in a cyanide bath containing turkey-red oil were remarkably good. The plated surface was found to follow the original one very closely indeed and was, in general, a little smoother. Many more experiments have now been carried out using cyanide baths with turkey-red oil and carbon disulphide brightening agents in an attempt to produce a plated surface smoother than the original surface.

(5.2) Copper

Bright-copper plating shows a great improvement over acid or cyanide copper plating with regard to smoothness of deposit and general levelling properties.[24] Bright-copper deposits are usually obtained by adding small quantities of colloidal material[25] (naphthol, gelatine, etc.) to either an acid or a cyanide bath. This results in an increase of lustre of the surface with a decrease in the grain size. Certain organic contaminants may cause dullness of the finish, but these may be eliminated by filtering or employing activated carbon.

A short length of precision-drawn brass guide has been plated by the use of an Efco-Udylite bright acid-copper solution for the purpose of examining the levelling properties of the plate. This bath is reported to give a high degree of lustre when carefully controlled—higher, in fact, than that obtained with a bright-silver solution. The procedure was as follows:

(1) Chemically polish for 1½min at 75°C in the same solution as used for the specimens described in Section 4.

(2) Rinse in cold water.
(3) Rochelle-copper flash for ½min at a current density of 50A/ft.^{2}
(4) Rinse in cold water.
(5) Dip in 10% solution of sulphuric acid.
(6) Rinse in cold water.
(7) Bright-acid copper-plate for 20min at a current density of 55A/ft^2 to a thickness of about 0·0008in.

Chemical polishing is used to produce a good surface to which electrodeposits adhere. Although this is no substitute for mechanical polishing for this purpose, it may usefully supplement it.[26]

(6) ELECTROFORMED GUIDES

Several sections of rectangular copper waveguide, for frequencies around 10 000Mc/s, electroformed commercially in different baths and on several types of mandrel, have been examined to determine which manufacturing technique gives the best internal surface.

First, two samples were electroformed on expendable mandrels of tin-lead-bismuth alloy, one (a 90° H bend) in an acid copper bath and the other (a magic-T junction) in a cyanide bath.

The effect of various types of stainless-steel mandrel on the internal roughness was then determined. A sample was electroformed in a cyanide bath and had a composite rectangular mandrel fabricated from stainless-steel plates with fusible-alloy plugs at the corners to hold them together. The mandrel was removed by melting the alloy. A second sample was electroformed in an acid bath on a permanent stainless-steel mandrel, which was withdrawn by applying a separating force from a screw mechanism. A third sample was produced on a similar type of mandrel, withdrawal in this instance being assisted by heating.

(7) SURFACE-ROUGHNESS MEASUREMENTS

(7.1) Procedure

There are many different ways of evaluating the roughness of a metal surface. Instruments having a stylus which is traversed horizontally across the surface may be used,[27,28,29,30] and numerous optical methods have been devised, mostly employing multiple-beam microinterferometers.[31,32,33,34] There are several objections to the use of both these types of instrument for recording the surface roughness of waveguide specimens. First, although the stylus type gives an indication of the centre-line average value of the surface and the peak roughness, it is not possible to measure the ratio of the actual length of a surface to its ideal length owing to unequal horizontal and vertical magnifications. This is the important information required for the present studies. In addition, the horizontal traverse of the stylus is only about 0·025in, and with deep scratches and a slightly worn stylus the stylus penetration may be only 25% of the full depth,[34] and false readings will be given. The optical methods obviate this disadvantage, but they, too, only give an indication of the roughness of a selected part of the sample and do not measure the ratio of the actual to the ideal length of the surface.

Consequently, it was decided to examine the surfaces by polishing and microscopic inspection in a manner similar to that successfully used on a previous investigation.[11] The method consists briefly in plating a sample with a metal contrasting in colour, mounting a section in a plastic mould, polishing first with various grades of abrasive paper and then with metal polish, and finally examining the surfaces with a microscope. It is desirable, however, to obtain definite values for the various surface-roughness factors, so, instead of taking measurements from typical photomicrographs of the section (using a magnification of about 1 000) and making statistical calculations, the ground-glass screen of the projection microscope was replaced by clear glass and tracing paper. In this way the whole surface contour of one side of a waveguide section can be traced, and accurate measurements can then be made of the length of the contour for comparison with the corresponding perfectly smooth surface. Values for the surface-roughness factors K_{T1}, K_{T2} and K_p may thus be obtained. In the case of plated surfaces the corresponding base-metal-surface tracings can also be obtained, thus giving a true measure of the levelling, if any, produced by the plating. Such measurements need to be limited to a small number of cross-sections, but experience has shown repeatedly that different sections from the same length of waveguide give surprisingly similar results. The main disadvantage of this method is that it is laborious; an enormous amount of energy and considerable time are necessary to obtain the required data. This is justifiable, however, since the results obtained by the process are very consistent and agree well with the values anticipated from the ratio between the measured and the calculated attenuation.[11]

Both skill and experience are required in polishing to produce a relatively scratch-free surface for examination. A different technique involving electropolishing has therefore been adopted recently which enables unskilled workers to produce a surface entirely free from scratches. With this process the protective plating need not be contrasting in colour, but it must be a metal which will electropolish in the same bath as the sample, i.e. copper plate on copper or brass. After the usual mounting and preliminary polishing with emery paper, the specimen is electropolished in a suitable bath. This produces an undistorted stress-free surface and eliminates the possibility of an error in the results due to masking of the surface by scratches. The method also saves an enormous amount of time and effort compared with the earlier method, since it is the final hand-polishing stages which require so much skill and patience. If the plated metal is the same as that of the guide, the sample is etched slightly during electropolishing by reducing the voltage, in order to make the junction of the original surface and the plating more obvious. The choice of electropolishing bath depends on the metal to be polished. Unfortunately, this new technique cannot be used successfully on silver-plated sections because silver will not polish in the baths used for copper and brass.

(7.2) Results

The measured surface-roughness factors for the various types of waveguide sample examined are given in Table 1.

Eqn. (13) gives the ratio of the calculated attenuation which takes account of surface roughness to that of a guide with perfectly flat surfaces, for the H_{01} mode. Values of this attenuation ratio have been calculated, using the experimental figures for the three roughness coefficients, for this commonly used mode, in an air-filled rectangular waveguide with internal dimensions 1in × 0·5in. The frequency has been taken as 10 000Mc/s. These attenuation ratios are also tabulated in Table 1, together with information previously determined for other types of 3cm waveguide surface, so that the performances of the various surfaces can be easily compared.

The main deformities of the surfaces from transverse sections of precision-drawn 3cm copper waveguides are a considerable number of "draw lines." They range in size from 0·001in deep by 0·0005in wide in the long sides to 0·0005in deep by 0·001in wide in the short sides. On the whole, the short sides are better than the long sides. The photomicrograph in Fig. 1 shows a typical transverse section taken from a short side. It is of interest to note that previous measurements[11] on a sample of normal

Table 1

ATTENUATION RATIOS AND SURFACE-ROUGHNESS FACTORS FOR VARIOUS TYPES OF WAVEGUIDE

Type of waveguide	K_{T1}	K_{T2}	K_p	α'_H/α_H
Electropolished copper	1·002	1·001	1·004	1·002
Electropolished brass	1·004	1·006	1·005	1·004
Normal-drawn brass*	—	—	—	1·060–1·074
Normal-drawn copper*	—	—	—	1·102–1·120
Precision-drawn copper	1·020	1·015	1·006	1·012
Precision-drawn copper*†	—	—	—	1·034
Precision-drawn brass	1·002	1·001	1·001	1·001
Precision-drawn brass*†	—	—	—	1·014
Acid-copper electroplate	1·625 max	1·625 max	—	—
Bright-copper electroplate	1·002	1·002	1·002	1·001
Silver-plate from cyanide bath*	—	—	—	1·580
Bright-silver electroplate	1·012	1·004	1·006	1·007
Chemically polished copper	1·002	1·005	1·005	1·003
Chemically polished brass	1·003	1·002	1·004	1·003
Electroformed copper on expendable alloy mandrel, Sample 1	1·018	1·025	1·013	1·017
Electroformed copper on expendable alloy mandrel, Sample 2	1·036	1·067	1·049	1·051
Electroformed copper on stainless-steel semi-permanent mandrel	1·010	1·021	1·001	1·008
Electroformed copper on stainless-steel permanent mandrel, Sample 1	1·006	1·005	1·003	1·003
Electroformed copper on stainless-steel permanent mandrel, Sample 2	1·001	1·002	1·001	1·001

* These values have been determined experimentally from attenuation measurements.
† The measured values of attenuation cannot be compared with the estimated attenuation for similar types of guide because the roughness measurements were carried out on guides of slightly better quality.

commercial brass waveguide showed that irregularities with magnitudes up to about 400 microinches occurred and that the ratio of the length of the actual surface to that of an ideally smooth one was larger for the short sides than for the long ones.

The term "microprofile"[35] is used throughout the following descriptions to mean a roughness which can easily be seen, but not measured, at a magnification of 1 000, and which is therefore comparable with the skin depth. In this respect the copper tube is smooth, whilst the general undulation is still present.

The longitudinal sections of drawn guides are not nearly so rough as the transverse sections, as pointed out in a previous paper.[11] There is a general absence of draw lines on these particular sections, the only irregularities being very shallow grooves about 0·001 in long by 0·0003 in deep, although there are occasional places where the microprofile is rough.

Precision-drawn brass waveguide is extremely good. It has already been noted[11] that normal commercial brass guides have a better internal finish than the corresponding copper ones. This is also true of precision-drawn tubes. The transverse-section contours are quite flat, usually showing only gentle undulations. Occasionally, there are more severe changes in the level of the surface. A few draw lines, about 0·0003 in deep,

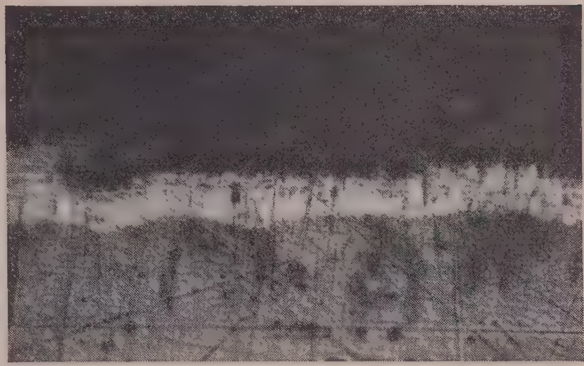


Fig. 1.—Typical photomicrograph of precision-drawn copper waveguide surface from 1 in × ½ in guide.

Section perpendicular to waveguide axis; magnification 500; copper in the lower part of the photograph.

much smaller and less numerous than for the corresponding copper guide, are in evidence, together with one or two shallow hollows. As for the copper guide, the microprofile is smooth. The short sides are smoother than the long ones. The microscopically smooth longitudinal-section surfaces only undulate very gently and there are no grooves or elevations. This brass guide is the best of all the drawn tubes examined and leaves little, if anything, to be desired as far as its internal roughness is concerned.

The surface roughness of some precision-drawn, 8 mm waveguide was examined to compare it with the 3 cm samples made by the same manufacturer. It was found that the guide is the roughest of all the precision-drawn guides tested, contrary to expectation from previous observations on several sizes of tube.[11] Surfaces from transverse sections showed severe undulations and elevations, some greater than 0·001 in, although the microprofile is quite smooth. Longitudinal-section surfaces are equally as rough as those from transverse sections, and certain parts of them have microscopic roughness as well as large undulations. The values of K_{T1}, K_{T2} and K_p are 1·028, 1·049 and 1·043 respectively.

A considerable improvement in the surface finish of precision-drawn copper tube was obtained by chemical polishing, which produced a very smooth finish with no acute draw lines but only a very gentle general waviness, as can be seen from the typical photomicrograph in Fig. 2. The results in Table 1 show the

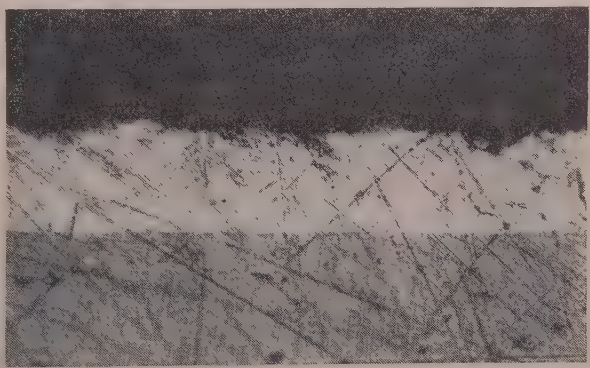


Fig. 2.—Typical photomicrograph of chemically polished copper waveguide surface from 1 in × ½ in guide.

Section perpendicular to waveguide axis; magnification 500; copper in the lower part of the photograph.

improvement over the copper tube described above due to polishing. Whereas K_p remains about the same, the transverse coefficient K_{T1} is reduced from 1·020 to 1·002 and K_{T2} from 1·015 to 1·005. Further, the copper tube selected for polishing was originally not so smooth as the tube discussed above because it had previously been determined[11] that its ratio of measured to calculated attenuation is 1·034.

The internal surfaces of chemically polished brass waveguide show only gentle undulations with a few sharp changes in level. No draw lines are present and there has been a microscopic polishing action. The most severe distortion of the surface takes the form of shallow hollows which appear on both the long and short sides, particularly near the corners. Whilst these can be smoothed out by chemical polishing, a good deal of metal would have to be removed.

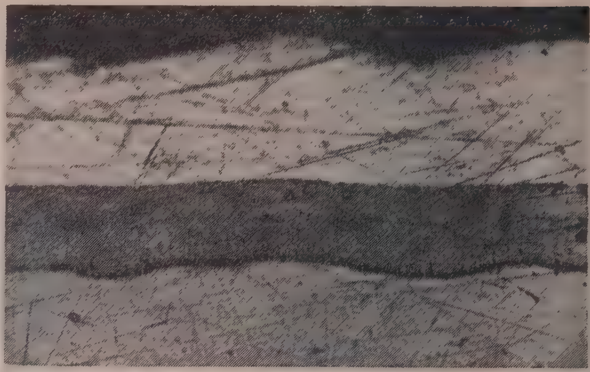
Although the roughness coefficients for this guide are slightly larger than those for the best-quality precision-drawn guide described above, they are much smaller than those for the original surfaces, which had a ratio of measured to calculated attenuation of 1·014.

The surface natures of the chemically polished brass and copper waveguides are similar, whereas the original copper guide is much rougher than the brass one. This suggests that chemical polishing is more effective on precision-drawn copper waveguides because of their original roughness, although it can be usefully employed on brass guides. The similarity in final finish obtained on both guides also emphasizes the fact that, although chemical polishing eliminates the microprofile, the general macroprofile still remains and would be difficult to eliminate without removing a considerable amount of metal from the surface.

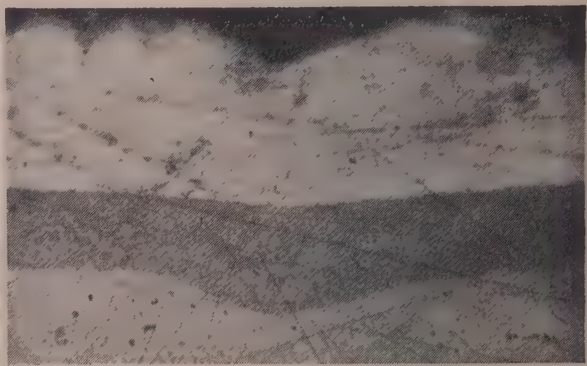
The results for bright-silver-plated surfaces are somewhat disappointing. In all cases the internal surface finish is considerably better when a bright bath rather than an ordinary cyanide bath is used. In general, the silver-plating faithfully reproduces any irregularity of the base metal. There is very little levelling of the base, however, and there are some cases where the silver has a rougher surface than the original.

Bright-copper plating appears to be smoother than the bright-silver deposits and has greater levelling properties (Fig. 3). Since the sample was plated on the outside, with no internal anode, the brass base surface is quite smooth, and it is therefore not surprising that the final copper surface is good. However, there is no exaggeration of the base irregularities as with cyanide and acid-copper plating.[12] Any irregularity of the base is either partially smoothed out or faithfully reproduced. All the roughness coefficients were found to be 1·001, the corresponding coefficients of the base all being 1·002. These figures give some indication of the levelling action. It cannot be assumed, of course, that the plating would be the same if the guide were internally plated, since the base metal is not so rough and the usual difficulties of throwing power and internal anodes are not encountered when the plating is outside the tube.

The surfaces of the longer sides of the transverse sections of electropolished copper guides are very smooth but undulate slightly, more especially towards the centre of the side. All microscopic roughness has disappeared owing to electropolishing, and there is no evidence of a microprofile even when the magnification is increased to as much as 4 000×, but the macroprofile remains, as was expected. There are no draw lines present and the only acute distortion takes the form of small pits in the surface, approximately 0·0001 in deep, as can be seen by the typical photomicrograph in Fig. 4. These are probably due to inclusions of copper oxide. The short sides appear to be much flatter, still retaining the same smoothness. The pits are also shallower and not so numerous.



(a)



(b)

Fig. 3.—Typical photomicrograph of copper-electroplated brass waveguide surface from 1 in × ½ in guide.

(a) Section perpendicular to waveguide axis.
(b) Section parallel to waveguide axis.
Magnification 500; brass in the lower parts of the photographs, bright-copper plating in the dark bands and the protective coating of silver at the tops.

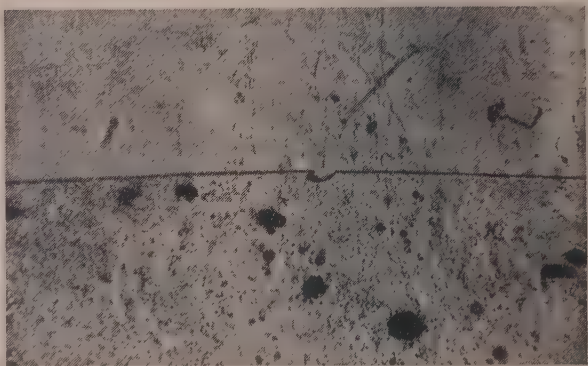


Fig. 4.—Typical photomicrograph of electropolished precision-drawn copper waveguide surface from 1 in × ½ in guide.

Section perpendicular to waveguide axis; magnification 500; waveguide in the lower part of the photograph.

A longitudinal section taken from a long side shows much the same characteristics as the long side of the transverse section. The surface is still very smooth but the general undulation is slightly worse and the pits are still present. Electrical tests performed previously on a long specimen of the same unpolished

guide indicate that the ratio of an actual length of surface to an ideal length is about 1·034.[11] It is evident, therefore, that a considerable improvement in the surface texture has been brought about by electrolytic polishing.

Electropolished brass waveguide surfaces also show considerable improvement over the original surfaces. As is characteristic of the electropolishing process, the microprofile is eliminated, but large undulations remain on the transverse sections. These undulations are worse than those for the corresponding copper guide, and include some sharp elevations and a number of severe changes in level. The topography of the surfaces from a longitudinal section is very similar to that from the transverse section. As with all sections of this specimen, there are no small pits in the surface such as were characteristic of the copper guide, thus indicating the absence of any included oxide. The original guide, which is smoother than the precision-drawn copper tube, had values of K_{T1} and K_{T2} of the order of 1·014. Thus, although electropolishing has produced a considerable improvement over the original surface, the technique is not so effective for brass as for copper guides.

The surface finish of the electroformed guides examined is dependent to a large extent on the type of mandrel employed and on the method of separation of the guide from the mandrel. A transverse section of the electroformed bend mentioned above shows that it has a very poor internal surface; it is quite undulous and microscopically rough in places. As is to be expected with guides electroformed on this type of fusible mandrel, the roughness of the short sides and of a section parallel to the guide axis is of the same order as that of a transverse section. The second sample made on an expendable mandrel has a much worse internal topography than the first, the microscopic roughness being much more general. The difference in internal roughness between the two samples may be attributed to a probable difference in mandrel preparation. A further reason for the discrepancy is the varying methods employed for the separation of the mandrel from the component. One method used is to melt out the alloy and finally to clean the specimen by blowing steam through it. Alternatively, sodium hydride may be used to dissolve the mandrel.

Transverse sections from the sample formed on the composite rectangular mandrel mentioned show that the surface is extremely flat and smooth over the majority of both the long and short sides. There are large cavities at the corners, however, caused by the fusible-alloy plugs, which contribute very greatly to the relatively high value of the roughness coefficients. Although the second sample does not suffer from this defect, it shows numerous scratch marks due to the withdrawal process. The resulting interior finish of the third sample is better than that of any other type of waveguide, whether electroformed, drawn, electropolished, electroplated or chemically polished. There is a complete absence of scratches due to the withdrawal technique, and the surface is almost perfectly flat and very smooth.

(8) CONCLUSIONS

The internal surface finish of the best precision-drawn waveguide is remarkably good. For this type of tube the increase in the attenuation due to imperfections in the internal finish is less than 2%, which is negligible for most applications. As expected from earlier work,[11] brass guides are much smoother than the corresponding copper ones. This is probably due to the softer nature of the copper and its greater liability to scratching, or to the higher friction between copper and the drawing dies. For higher-frequency working, the waveguide dimensions decrease, and the roughness of an 8 mm sample was unexpectedly found to be considerably greater. The theoretical increase of attenuation due to the observed roughness is about 4% at a frequency of 37 000 Mc/s, for the H_{01} wave.

Chemical polishing or electropolishing can be employed to produce an even smoother surface than that of a precision-drawn guide. It is doubtful, however, whether such additional processing is worth while on a commercial scale, since the surface finish of precision-drawn guides is quite satisfactory for all normal purposes. The processes may find a useful application in the preparation of special components such as cavities, which must have a high degree of surface finish to give a high Q-factor. Chemical-polishing solutions may find a more general application than electropolishing since they can be used economically and the necessity for auxiliary anodes when polishing complex components is obviated. It should be remembered that the optimum voltage in an electropolishing bath is very critical; higher and lower voltages cause pitting and etching, respectively.

The texture of the inside of electroformed waveguides is found to be dependent on the type of mandrel used and the method of separation. The results show that stainless-steel mandrels give much better surfaces than mandrels of other alloys. Then, from the roughness point of view, waveguide T-junctions and similar complex components should be electroformed on multiple stainless-steel mandrels rather than on fusible types.

Where possible, also, mandrels relying on differential expansion for separation should be used in preference to composite types with fusible-alloy plugs. With a well-manufactured stainless-steel mandrel it is possible to electroform copper guide which has a better surface finish than any other type of waveguide.

Although the levelling produced by the different plated metals is not the same for them all, a bright-plated surface always shows a considerable improvement over the ordinary plated surfaces examined previously.[12] It is thought that the plating examined during the present investigations is the best that can be done with ordinary plating techniques. Bright-silver plating is unreliable; sometimes it exhibits no levelling properties and the finished surface may be worse than the original one. Thus it is doubtful whether bright-silver plating is worth while as a means of further reducing attenuation, because the increase in attenuation due to the rougher surface may well be greater than the decrease due to the higher conductivity of the silver. If drawn brass guide is bright-copper plated, however, an improvement is possible because of the increased conductivity and good levelling properties of the copper.

It is evident from eqns. (10) and (13) that the ratio α'_H/α_H depends on the order of the mode as well as on the roughness, because of changing values of n and m. Hence, all wave modes will not give the ratios of α'_H/α_H calculated for the H_{01} mode and given in Table 1. It is also interesting to note from eqn. (12) that the ratio α'_E/α_E is identical with K_p for all modes. It follows that the measured attenuation for an E wave will not be very different from that calculated on the assumption of perfectly flat surfaces, even for rather poor drawn tube where grooves along the length may produce high values of K_{T1} and K_{T2}.

So far, there has been no reliable way of estimating the performances of waveguide surfaces without making careful measurements on long samples of guide with the aid of elaborate test equipment. It is thought that manufacturers will find the method described in the paper useful for making quick and accurate examinations of waveguides.

(9) ACKNOWLEDGMENTS

The work recorded in the paper has been carried out in the Department of Electrical Engineering at the University of Sheffield. The authors wish to thank Mr. O. I. Butler for facilities afforded in the laboratories of this Department and

for his constant encouragement and advice. They are also indebted to Mr. H. T. Bache of the Electro-Chemical Engineering Co. Ltd., to Evered and Co. Ltd. and to Mr. E. W. Wilson for supplying waveguide samples for examination, and they are grateful to Dr. R. W. K. Honeycombe for his helpful criticisms and valuable suggestions.

(10) REFERENCES

(1) Vivian, A. C.: "The Superficial Conductivity of Metallic Conductors for Waveguides," Telecommunications Research Establishment Technical Note No. 45, 1949.

(2) Morgan, S. P.: "Effect of Surface Roughness on Eddy Current Losses at Microwave Frequencies, *Journal of Applied Physics*, 1949, **20**, p. 352.

(3) Maxwell, E.: "Conductivity of Metallic Surfaces at Microwave Frequencies," *ibid.*, 1947, **18**, p. 629.

(4) Milner, C. J., and Clayton, R. B.: "Resistance Comparison at 3 300 Mc/s by a Novel Method," *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 1409.

(5) Beck, A. C., and Dawson, R. W.: "Conductivity Measurements at Microwave Frequencies," *Proceedings of the Institute of Radio Engineers*, 1950, **38**, p. 1181.

(6) Beck, A. C.: "Conductivity Measurements at Microwave Frequencies," *Bell Laboratories Record*, 1950, **28**, p. 438.

(7) Horner, F., Taylor, T. A., Dunsmuir, R., Lamb, J., and Jackson, W.: "Resonance Methods of Dielectric Measurement at Centimetric Wavelengths," *Journal I.E.E.*, 1946, **93**, Part III, p. 53.

(8) Miller, S. E., and Beck, A. C.: "Low-Loss Waveguide Transmission," *Proceedings of the Institute of Radio Engineers*, 1953, **41**, p. 348.

(9) Sims, G. D.: "The Influence of Bends and Ellipticity on the Attenuation and Propagation Characteristics of the H_{01} Circular Waveguide Mode," *Proceedings I.E.E.*, Monograph No. 58 R, January, 1953 (**100**, Part IV, p. 25).

(10) Chambers, R. G., and Pippard, A. B.: "The Effect of Method of Preparation on the High-Frequency Surface Resistance of Metals," Institute of Metals Monograph, No. 13, 1953, p. 281.

(11) Benson, F. A.: "Waveguide Attenuation and its Correlation with Surface Roughness," *Proceedings I.E.E.*, Paper No. 1467 R, March, 1953 (**100**, Part III, p. 85).

(12) Benson, F. A.: "Attenuation and Surface Roughness of Electroplated Waveguides," *Proceedings I.E.E.*, Paper No. 1518 R, July 1953 (**100**, Part III, p. 213).

(13) Kuhn, S.: "Calculation of Attenuation in Wave Guides," *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 663.

(14) Kemsley, D. S., and Tegart, W. J. McG.: "The Electrolytic Polishing of Metals in Research and Industry," Australian Council for Scientific and Industrial Research Physical Metallurgy Report No. 7, Melbourne, 1948.

(15) Jacquet, P. A.: "Principles and Scientific Applications of the Electropolishing of Metals," Proceedings of the Third International Conference of Electrodeposition, September, 1947.

(16) Elmore, W. C.: "Electrolytic Polishing," *Journal of Applied Physics*, 1939, **10**, p. 724, and **11**, p. 797.

(17) Pinner, R.: "Theory of Electrolytic Polishing," *Electroplating*, 1953, **6**, p. 444.

(18) Halut, R. E.: "Technical Applications of Electrolytic Polishing," Proceedings of the Third International Conference of Electrodeposition, September, 1947.

(19) Pinner, R.: "Theory and Practice of Chemical Polishing, Part I," *Electroplating*, 1953, **6**, p. 360.

(20) Pinner, R.: "Theory and Practice of Chemical Polishing, Part 2," *ibid.*, 1953, **6**, p. 401.

(21) Sterenguel, J.: "Le polissage chimique de l'aluminium et de ses alliages," *Revue de l'aluminium*, 1953, **30**, p. 261.

(22) Meyer, W. R.: United States Patent No. 2088429, 1948.

(23) Powell, E. C., and Davies, R.: United States Patent No. 2113517, 1948.

(24) Raub, E.: "Glänzende galvarische Metallniederschläge," *Metalloberfläche*, Ausgabe B, March, 1953, **7**, p. A33.

(25) Wernick, S.: "Electrolytic Polishing and Bright Plating of Metals" (Alvin Redman, 1948).

(26) Wullhorst, B.: "Anodisches und Chemisches Glänzen als Vorbereitung für die nachfolgende Galvarisierung," *Metalloberfläche*, Ausgabe A, February, 1953, **7**, p. A28.

(27) Reason, R. E., Hopkins, M. R., and Garrod, R. I.: "Report on the Measurement of Surface Finish by Stylus Methods" (Taylor, Taylor and Hobson Ltd., 1944).

(28) Wright-Baker, H.: "Modern Workshop Technology," Part II, Chapter 14 (Cleaver-Hume, 1950).

(29) Schmidt, W.: "Oberflächenmessung mit Tastschnittgeräten,' *Metalloberfläche*, Ausgabe A, March, 1953, **7**, p. A33.

(30) British Standard No. 1134, 1950: "The Assessment of Surface Texture, Centre-Line-Average-Height Method," p. 52.

(31) Thieisch, H.: "Visual and Optical Evaluation of Metal Surfaces," *Metal Finishing*, May, 1951, **49**, p. 54.

(32) Stray, W., and Ogburn, G.: "Measurement of Surface Roughness with Microinterferometer," Proceedings of the American Electroplaters Society 37th Annual Convention, 1950, p. 125.

(33) Sugg, R. E.: "An Interferometer for Examining Polished Surfaces," *Mechanical Engineering*, 1953, **75**, p. 629.

(34) Timms, C., and Scoles, C. A.: "Surface Finish Measurements: Comparison of Stylus and Interference Methods," *Metal Treatment*, 1951, **18**, p. 450.

(35) Jacquet, P., and Rocquet, P.: "Application du polissage électrolytique à l'examen micrographique du fers et des aciers," *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, 1939, **208**, p. 1012.

(11) APPENDIX

(11.1) Modification of Kuhn's Attenuation Formulae to include the Effects of Surface Roughness

The general expression derived by Kuhn[13] for the attenuation in rectangular waveguides is

$$\alpha = \left[\frac{c}{\sigma}\frac{\mu_1}{\mu}\left(\frac{\epsilon}{\mu}\right)^{1/2}\right]^{1/2} \frac{(\mu/\epsilon)^{1/2}}{8ab(\lambda_e)^{1/2}k_s W'} \Sigma \int_s |H_w|^2 ds$$

where k_s is a numerical factor having a different value for different wave modes, W' is a quantity Kuhn calls the characteristic density of energy, and the integration is taken along the whole periphery, s, of the guide cross-section. In the case of an H_{nm} mode, $k_s = \frac{1}{2}$ if $n = 0$ and $k_s = \frac{1}{4}$ if $n \neq 0$. For an E_{nm} mode $k_s = \frac{1}{4}$.

The sums of the integrals in eqn. (14) are evaluated below for H_{nm} and E_{nm} waves. In making these calculations the field equations given by Kuhn in right-handed Cartesian co-ordinates are used. The origin of the system is placed in one of the corners of the guide cross-section so that the width b falls in the positive direction of the z axis, the height a in the positive direction of the y axis and the axis of the guide extends in the direction of the x axis.

(11.2) Attenuation of an H_{nm} Wave in a Rectangular Waveguide

The top and bottom walls and the side walls of the guide are considered separately.

At the top and bottom walls (i.e. for $y = 0$ and $y = a$) only the components of H_x and H_z of the magnetic field H_w are found. They vary in the z direction with $0 \leqslant z \leqslant b$. In the case of perfectly smooth surfaces these walls contribute an amount T_1 towards the summation in eqn. (14) such that

$$T_1 = 2\left(\int_0^b |H_x|^2 dz + \int_0^b |H_z|^2 dz\right) \quad . \quad . \quad . \quad (15)$$

When the surfaces are rough, since the irregularities are, in general, much greater than the skin depth, expression (15) becomes

$$T_1 = 2\left(\int_0^b K_{T1}|H_x|^2 dz + \int_0^b K_p|H_z|^2 dz\right) \quad . \quad . \quad (16)$$

This is so because of the increase in the effective resistance of the associated current paths due to roughness. K_{T1} and K_p are roughness factors already defined.

Using the formulae for H_x and H_z given by Kuhn, it is found that

$$T_1 = 2W'Z_g b\left(\frac{K_{T1}}{Z_{cr}^2} + \frac{K_p}{Z_g^2}\frac{Z_{cr}^2}{Z_b^2}\right) \quad . \quad . \quad . \quad (17)$$

Similarly, for the side walls, where $z = 0$ or $z = b$, the components of H_x and H_y are generally present and they vary in the y direction with $0 \leqslant y \leqslant a$. The sum of the integrals to be used in eqn. (14) is now different in the cases in which $n = 0$ and $n \neq 0$.

When $n \neq 0$ the contribution made by the side walls towards the summation is T_2 such that

$$T_2 = 2\left(\int_0^a K_{T2}|H_x|^2 dy + \int_0^a K_p|H_y|^2 dy\right) \quad . \quad . \quad (18)$$

The same value of K_p has been used for the top and bottom walls and the side walls in eqns. (16) and (18) because measurements show that such an assumption is quite justified.

Using the formulae for H_x and H_y given by Kuhn in expression (18), it is found that

$$T_2 = 2W'Z_g a\left(\frac{K_{T2}}{Z_{cr}^2} + \frac{K_p}{Z_g^2}\frac{Z_{cr}^2}{Z_a^2}\right) \quad . \quad . \quad . \quad (19)$$

If $n = 0$, $H_y = 0$ and

$$T_2 = 4W'Z_g a\frac{K_{T2}}{Z_{cr}^2} \quad . \quad . \quad . \quad . \quad . \quad (20)$$

In the above equations Z_g is the specific impedance of the guide, which is the ratio of the transverse components of the electric and magnetic fields responsible for propagation along the guide, and Z_{cr} is the transverse impedance of the guide, which is the ratio of those components of fields which are responsible for transverse oscillations. Z_a and Z_b are two kinds of transverse impedance introduced by Kuhn.[13]

In the case of an H-type wave,

$$Z_g = \left(\frac{\mu}{\epsilon}\right)^{1/2}\frac{\lambda_g}{\lambda_e} \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

$$Z_{cr} = \left(\frac{\mu}{\epsilon}\right)^{1/2}\frac{\lambda_{cr}}{\lambda_e} \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

$$Z_a = \left(\frac{\mu}{\epsilon}\right)^{1/2}\frac{2a}{n\lambda_e} \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

and

$$Z_b = \left(\frac{\mu}{\epsilon}\right)^{1/2}\frac{2b}{m\lambda_e} \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

From eqns. (17) and (19) the total sum of the integrals for the case in which $n \neq 0$ is

$$T_1 + T_2 = 2W'\frac{\lambda_g}{\lambda_e}\left(\frac{\epsilon}{\mu}\right)^{1/2} aK_r\left(\frac{\lambda_e^2}{\lambda_{cr}^2} + K_t\right) \quad . \quad . \quad (25)$$

Substitution of expression (25) in eqn. (14) gives eqn. (2).

From eqns. (17) and (20) the total sum of the integrals for the case in which $n = 0$ is

$$T_1 + T_2 = 4W'\left(\frac{\epsilon}{\mu}\right)^{1/2}\frac{\lambda_g}{\lambda_e}a\left[K_{T2}\frac{\lambda_e^2}{\lambda_{cr}^2} + K_{T1}\frac{b}{2a}\frac{\lambda_e^2}{\lambda_{cr}^2} + K_p\frac{b}{2a}\left(1 - \frac{\lambda_e^2}{\lambda_{cr}^2}\right)\right] \quad . \quad . \quad . \quad . \quad (26)$$

Substitution of expression (26) in eqn. (14) gives eqn. (1).

(11.3) Attenuation of an E_{nm} Wave in a Rectangular Waveguide

For the case of an E_{nm} wave the calculations involved in using eqn. (14) are made on similar lines to those for an H_{nm} wave, but are simpler owing to the fact that only the transverse magnetic field is encountered. The value of n cannot be equal to zero in this case.

After integration, the following values for the sums of the integrals are obtained:

For the top and bottom walls

$$2\frac{W'}{Z_g}\frac{Z_a^2}{Z_{cr}^2}bK_p \quad . \quad . \quad . \quad . \quad . \quad . \quad (27)$$

For both side walls

$$2\frac{W'}{Z_g}\frac{Z_b^2}{Z_{cr}^2}aK_p \quad . \quad . \quad . \quad . \quad . \quad . \quad (28)$$

For an E-type wave

$$Z_g = \left(\frac{\mu}{\epsilon}\right)^{1/2}\frac{\lambda_e}{\lambda_g} \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

$$Z_{cr} = \left(\frac{\mu}{\epsilon}\right)^{1/2}\frac{\lambda_e}{\lambda_{cr}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (30)$$

$$Z_a = \left(\frac{\mu}{\epsilon}\right)^{1/2}\left(\frac{n\lambda_e}{2a}\right) \quad . \quad . \quad . \quad . \quad . \quad (31)$$

and

$$Z_b = \left(\frac{\mu}{\epsilon}\right)^{1/2}\left(\frac{m\lambda_e}{2b}\right) \quad . \quad . \quad . \quad . \quad . \quad (32)$$

By summation of eqns. (27) and (28) and substitution in eqn. (14), expression (7) is obtained.

DISCUSSION ON

"RECEPTION OF AN F.M. SIGNAL IN THE PRESENCE OF A STRONGER SIGNAL IN THE SAME FREQUENCY BAND, AND OTHER ASSOCIATED RESULTS"*

Mr. E. I. Bosch (*United States: communicated*): The paper contains a rather clever scheme (Section 9) of putting two intelligences in one frequency channel without requiring increased sharpness of filtering. It may facilitate comprehension to point out that this scheme is equivalent to frequency-modulating the r.f. carrier with intelligence A, and then single-sideband-modulating the resultant signal with a subcarrier which is frequency-modulated with intelligence B. Strangely, the paper makes no effort to reconcile the implied saving in bandwidth with the principles of information theory. Section 11 (Comparative Analysis of Bandwidth Required) is quite inconclusive, since it does not consider signal/noise ratios, transmitted powers, or information bandwidths. Furthermore, I was unable to find any substantiation of the last sentence in Section 11.1, which states that in an f.m. system whose modulation consists of an audio signal and a subcarrier, "the total bandwidth is equal to that taken up on account of the subcarrier plus $2f_1$," where f_1 is the carrier-frequency deviation due to the audio modulation alone.

The paper's methods of separating weak f.m. signals from strong ones appear to apply only when the frequency difference between the two carriers is considerably greater than the modulation frequencies on them. This condition is contained in Section 4.1, where it is suggested that ω_2, the instantaneous frequency difference between the carriers, be measured and added to (or subtracted from) ω_1, the instantaneous frequency of the stronger carrier, in order to obtain the frequency of the weaker one. Such a method requires that ω_2 should make a sufficient number of cycles to be measured, in a time in which the two carrier frequencies remain substantially constant. Again, in Section 4.2, the symbols ω_1 and ω_2 are used both in eqn. (7) and eqn. (8), even though those equations refer to two different times ($\theta = 0$ and $\theta = \pi$). This would not be permissible if it were not assumed that ω_1 and ω_2 change but very little in the interval from $\theta = 0$ to $\theta = \pi$ (a half-cycle of ω_2). That is, it is assumed that the modulation frequency is much smaller than the frequency difference between the two carriers. There arises, therefore, the question whether the method is—without major modification—applicable to co-channel interference, where the frequency difference between the carriers is usually of the same order of magnitude as the higher modulation frequencies.

Two minor points: Section 10.2 states that the amplitude variation at the input to the limiter can be increased by using "an automatic gain control operating, preferably on sharp-cut-off tubes, to cut off more of the lower part of the r.f. oscillation as the signal increases in intensity without distorting the shape of the envelope." This is incorrect; such a circuit will increase the relative amplitude variation, but decrease the absolute variation.

Section 10.3 states that, instead of using a simple discriminator "it may be more satisfactory to use a circuit that establishes a point on the waveform at each cycle of frequency." Is there no danger that points thus established on waveforms might blur, or, under adverse conditions, even fade completely?

* WILMOTTE, R. M.: Paper No. 1603 R, March, 1954 (see **101**, Part III, p. 69).

Mr. R. M. Wilmotte (*in reply*): In Section 9 it is described how intelligences are combined in one frequency channel and may be separated without requiring increased sharpness or filtering; the equivalent operation given by Mr. Bosch is not quite the same, and, unless I have misunderstood that equivalence, it does not lead to an identical mathematical expression for the signals.

Mr. Bosch points out that the paper does not analyse the system in terms of the modern information theory. That failure I do not think is strange. It seems somewhat strange to me that it is thought to be strange and the reason is simple. Modern information theory has been applied largely to problems of evaluating the signal/noise ratio, usually eliminating considerations of other forms of interference. Information theory has developed in that manner largely because the signal/noise ratio can be handled with relatively simple mathematics. When the problems become too complicated, mathematicians tend to leave them to their successors. The result has been—and this I give as my personal criticism of the existence of an undesirable trend of the application of information theory—an over-emphasis of the importance of signal/noise ratio. In this particular case, that which probably would prove to be the limiting factor in the operation of the system is not noise but the distortion due to multi-path signals. In any case, an analysis of these f.m. signals with their associated complex and non-linear circuits, in terms of information theory, would probably lead to complex mathematics. Even if the theory were worked out and results obtained that could be considered accurate, it would give only a partial picture since the only interference considered would have been noise.

The paper discusses briefly the bandwidth required, and this was considered not from the point of view of signal/noise ratio, but for its effect on adjacent channels. It appears that at least part of the explanation of my analysis was not clear. The part that seems to be difficult was that which dealt with the bandwidth required in the usual f.m. system, in which a carrier is frequency modulated with signal A plus a subcarrier which is itself modulated with signal B. The bandwidth required for such a system is one that involves statistical evaluation of the time during which the carrier modulated with signal A is deviated in frequency from the mean frequency by a substantial amount. If this signal is voice or music, the maximum deviation will occur, to a large extent, at the low notes. A low frequency implies that the carrier for a considerable period of time corresponding to the periods around positive and negative maxima of the signal is deviated at or close to a 100% deviation. If we consider a subcarrier added to this signal, all the frequencies in that subcarrier are very high compared with the frequencies of this low note, so that from the point of view of the subcarrier, when that low note produces a wide deviation, say f_1, the main carrier may be considered as having been shifted by an amount f_1. The subcarrier is then applied to this deviated carrier. The first sideband of the subcarrier is at a frequency equal to that of the subcarrier, f_2, so that this sideband occurs at a frequency $(f_1 + f_2)$ away from the main carrier. A similar situation occurs when the frequency swing is negative. Thus the

total bandwidth is at least equal to $2(f_1 + f_2)$. It is incorrect to consider the bandwidth as though signal A and the modulated subcarrier were merely equivalent to signal A alone, but amplified to produce the same peak amplitude. The basic reason is that the frequency distribution in the two systems is not the same statistically. When a subcarrier is used, there is at all times a large amount of very-high frequencies, which is not true when signal A alone is used. This explanation was not clearly stated in the paper: I had, probably improperly, tried to state it in the space of one paragraph.

There is much work left undone in developing proper circuits for clearly separating a weak f.m. signal from a strong one. The paper merely described possible techniques. One method was to measure ω_1, the angular velocity of one signal, and ω_2, the angular velocity of the difference between the two signals. These two should be measured at the same instant and continuously. The fact that they are not so measured will produce some distortion, and the degree of distortion produced will depend a great deal on the system used. It is not true that it is necessary that ω_2 should cover "many cycles" in order to measure it. In fact, the methods referred to in the paper of Krumhansl and Goldberg, and those of Hoeppner, are able to provide adequate information on ω_2 within one cycle. The possibility may be explained from eqn. (1), where ω_1 and ω_2 are referred to as the rate of change of phase, rather than proportional to frequency.

It is also true that in eqns. (7) and (8) the two factors ω_1 and ω_2 are measured at different times so that that method will probably produce some distortion when $(\omega_1 - \omega_2)/2\pi$ lies within the intelligence frequency band. The suggestion by Mr. Bosch that many cycles are required for these measurements is not justified except from our experiments with a simple limiter-discriminator circuit. With it, substantial distortion was obtained, but it is to be expected that better circuits can be developed. Great improvement was obtained by using the circuits referred to above. A brief statement indicating the advantages of these circuits is made in Section 10.3.

As to how small a difference in the mean frequencies can be maintained without developing serious distortion is not known. It is believed, however, that with proper circuit technique it may be possible to operate when the two channels are very close together in frequency. That is indicated broadly in the last paragraph of Section 4.1. In support of that concept are the results obtained by Dr. Argimbau, who, with the use of a very-wide-band discriminator, has been able to obtain substantial distortionless reception when one f.m. signal was in the presence of another one only a fraction of a decibel lower in intensity. The distortion that occurs under those conditions is indicated in Fig. 2. It will be seen there that within one cycle of the difference of frequency between the two signals there is a tremendous degree of distortion. This distortion does not appear, however, in wideband f.m. reception when the Argimbau type of circuit is used because the waveform within one cycle is averaged accurately without distortion. When the frequency difference between the two signals is low and audible, some distortion should be heard. The fact that the fascinating experiments of Argimbau have produced such excellent results leads one to expect, or at least to hope, that similar results would be obtained in the case of the application of the systems described in the paper.

As regards Section 10.2, referring to a suggestion for increasing the relative variation in the waveform, there seems to be agreement with Mr. Bosch that the system will increase the relative amplitude. The absolute variation can be controlled by increasing the gain. Unless the relative amplitude is first increased, there is the danger of reducing the relative amplitude by overloading the tubes. It is my impression that the statement in the paper is correct, but possibly not developed adequately.

As regards the use of the Krumhansl and Goldberg, or Hoeppner, circuits (Section 10.3), I can only state that these have been tried in the laboratory and have proved a considerable improvement over previous circuits. I do not see why "points" thus established on the waveform might "blur or even fade completely."

The paper was written with the hope that it might stimulate engineers to a somewhat new approach to certain f.m. techniques. It is by no means a complete analysis of the potentialities of circuit technique that may prove applicable, but it is believed that the techniques may have valuable potentialities for certain applications.

DIGESTS OF INSTITUTION MONOGRAPHS

THE EFFECT OF SEVERE AMPLITUDE LIMITATION ON CERTAIN TYPES OF RANDOM SIGNAL: A CLUE TO THE INTELLIGIBILITY OF "INFINITELY" CLIPPED SPEECH

621.391 : 621.3.018.7 Monograph No. 111 R

J. M. C. DUKES, M.A., Associate Member.

(DIGEST *of a paper published in November*, 1954, *as an* INSTITUTION MONOGRAPH *and republished in March*, 1955, *in Part C of the* PROCEEDINGS.)

Fig. 1 shows the input/output characteristic of a peak-limiting amplifier. If the input signal of Fig. 2(*a*) has an average amplitude very much larger (e.g. 60 dB) than AO in Fig. 1, the output waveform will consist for practical purposes of a train of rectangular pulses as in Fig. 2(*b*). This process, sometimes referred to as "infinite" peak clipping has surprisingly little effect on the intelligibility of conversational speech. More specifically, tests[1–3] based on randomly selected one-syllable words show articulation scores of the order of 80–90%. An initial period of accommodation is required whilst the listener gets used to the noisy character of the sound, and moreover fatigue is likely to set in fairly rapidly if the test continues too long. Neither subsequent differentiation [Fig. 3(*c*)] nor integration have any significant effect on intelligibility, although in the latter case there is a welcome reduction in quantization noise.

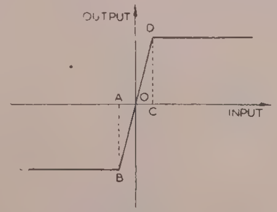


Fig. 1.—Amplifier characteristic.

Mr. Dukes is with Standard Telecommunication Laboratories, Ltd.

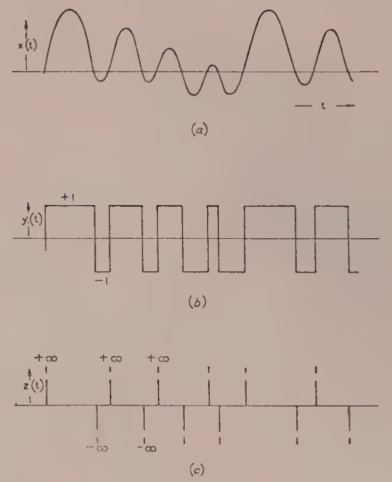


Fig. 2.—Effect of "infinite" peak clipping on a typical waveform.

(a) Original waveform.
(b) Clipped waveform.
(c) Clipped waveform with subsequent differentiation.

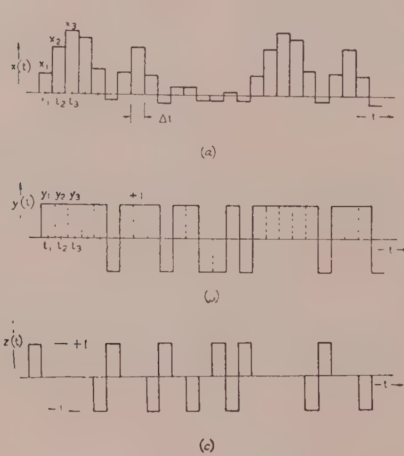


Fig. 3.—Time-quantized waveforms.

(a) Original pulse train [time-quantized version of Fig. 2(a)].
(b) Clipped pulse train.
(c) Differentiated clipped pulse train, with pulse reshaping.

In comparing Figs. 2(a) and 2(b) it is evident that the only characteristic retained is the relative position of the zero crossings of the signal. The conclusion that the greater part of the information-bearing content of speech must be carried by the zero crossings is inevitable, but it is not entirely satisfactory in view of the almost universally accepted concept of the ear as a short-time spectrum analyser. The object of the paper, therefore, is to examine the relationship between the signals before and after clipping on a purely spectral basis.

For mathematical convenience the continuous waveforms of Fig. 2 are time-quantized in the manner illustrated in Fig. 3. This is permissible in view of the well-known proposition relating the amount of detail in a signal to the bandwidth of the circuits transmitting the signal. Provided Δt is sufficiently small, the time-quantized signal may be passed through a suitably shaped low-pass filter and the original signal thereby recovered without distortion. As regards the signal amplitudes, the analysis is equally applicable whether the range of possible amplitude levels be continuous or quantized in discrete levels.

The original pulse train of Fig. 3(a) is assumed to be generated by some form of stationary random process;[4] i.e. the statistics of the signal are invariant under any shift in the origin of time. No consideration is given to purely periodic waveforms, as these may be handled by simpler and more conventional techniques. It is necessary, however, to distinguish clearly between two classes of random signals, for which the following nomenclature has been adopted:

Totally random signals.—These are produced by a generating process such that the amplitudes of successive pulses are independently selected.

Partially constrained signals.—These are such that the probability of a given amplitude is partly dependent on the preceding sequence of amplitude levels.

A simple example of the first class is the sequence of numbers produced by the throwing of a dice. If the dice is "loaded" some numbers become more probable than others, but the process is otherwise unmodified. Into the second class come signals of a more complex character such as speech. However, provided that the quantities to be evaluated are invariant of the constraints between successive signal values, speech may be treated as though it were a totally random signal. This assumption is inherent in the analysis which follows.

The totally random signal may be defined completely by means of the one-dimensional probability distribution $p(x)$. No restriction is placed on the form of $p(x)$ except that the d.c. level of the original signal is zero,

i.e. $$\int_{-\infty}^{+\infty} xp(x)dx = 0$$

Two specific distributions for $p(x)$ are shown in Fig. 4, σ_x being the standard deviation (i.e. r.m.s. amplitude) in each case.

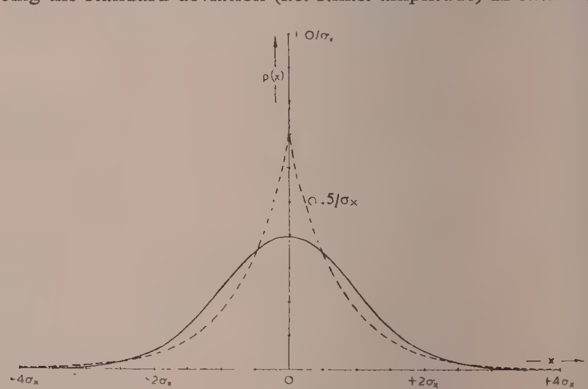


Fig. 4.—Probability distributions.

——— $p_1(x) = \frac{1}{\sigma_x\sqrt{(2\pi)}}\varepsilon^{-x^2/2\sigma_x^2}$

— — — $p_2(x) = \frac{1}{\sigma_x\sqrt{2}}\varepsilon^{-\sqrt{2}|x|/\sigma_x}$

The first is the Gaussian distribution, which may be used to a close approximation to represent the instantaneous amplitude distribution of unvoiced speech sounds.[5] The second, sometimes referred to as the exponential distribution, is a close approximation to the instantaneous amplitude distribution of voiced speech sounds.[5]

Now a random signal, like any other signal, may also be described by a complex spectral function $F_x(f)$. The infinitely clipped signal may likewise be defined by a function $F_y(f)$. These functions differ from those associated with purely periodic waveforms in that the phase component for each frequency can

only be specified statistically. In order to examine the relationship between $F_x(f)$ and $F_y(f)$, two criteria will be used, both based on the product $\overline{F_x(f)F_y(f)}$, where the bar indicates an average over all possible phase displacements $\phi_x(f) - \phi_y(f)$. Had the two random signals been independently generated this average product would tend to zero for all values of f. However, when $y(t)$ is derived from $x(t)$ by some process such as clipping, the value is likely to be other than zero, either positive or negative. Finally, averaging over all values of f, normalizing, and adjusting the delay τ to a value τ' so as to maximize the result, we obtain

$$\mu_{xy} = \frac{1}{\sigma_x \sigma_y} \left| \int_{-\infty}^{+\infty} \overline{F_x(f)F_y(f)} \varepsilon^{j2\pi f\tau'} df \right| \quad . \quad . \quad . \quad (1)$$

where μ_{xy} is the first coherence coefficient—a term adopted in the paper for want of a more suitable nomenclature.

It can be shown that the above expression for μ is numerically equal to

$$\text{Limit } T \to \infty \frac{1}{2\sigma_x \sigma_y T} \int_{-T}^{+T} x(t)y(t + \tau')dt \quad . \quad . \quad (2)$$

This will be recognized as the cross-correlation coefficient[4] of the two signals for the given delay τ'. In other words, the first coherence coefficient is equal to the maximum numerical value of the correlation function $\rho_{xy}(\tau)$.

A careful examination of eqn. (2), however, reveals one unsatisfactory aspect of μ as a definition of coherence. As a simple illustration, consider the result obtained for the input/output relationship of a passive all-pass filter having a non-linear phase characteristic. Evidently, in this case μ must of necessity be numerically less than unity because there is no single value of τ which will ensure that both $F_x(f)$ and $F_y(f)$ are in phase throughout the entire spectral range of both signals. On statistical grounds, however, these two signals might be considered as being coherent, and this viewpoint would be in accordance with the observation that the ear is relatively insensitive to phase distortion. For this reason it is advantageous to define a second coherence coefficient, totally independent of τ, such that

$$\nu_{xy} = \frac{1}{\sigma_x \sigma_y} \int_{-\infty}^{+\infty} \left| \overline{F_x(f)F_y(f)} \right| df \quad . \quad . \quad . \quad . \quad (3)$$

Applying these two criteria to the case of an infinitely clipped signal the following result is obtained:

$$\nu_{xy} = \mu_{xy} = A_x(1 + k)/\sigma_x k^{\frac{1}{2}} \quad . \quad . \quad . \quad . \quad (4)$$

where

$$A_x = \int_0^{\infty} xp(x)dx = -\int_{-\infty}^{0} xp(x)dx \quad . \quad . \quad . \quad (5)$$

$$k = \left[\int_{-\infty}^{0} p(x)dx\right] \bigg/ \left[\int_0^{+\infty} p(x)dx\right] \quad . \quad . \quad . \quad (6)$$

$$\sigma_x = \left[\int_{-\infty}^{+\infty} x^2 p(x)dx\right]^{\frac{1}{2}} \quad . \quad . \quad . \quad . \quad . \quad (7)$$

This gives the following numerical result for the two distributions of Fig. 4:

$\mu_{xy} = 0 \cdot 798$, for the Gaussian distribution,

and $\mu_{xy} = 0 \cdot 707$, for the exponential distribution.

These results are of special interest in that they have a striking numerical similarity to the articulation scores mentioned at the outset. Furthermore, it is found that, in practice, clipping is more destructive on the vowel sounds than the consonants. (It is of importance to note that the results are valid irrespective of the existence of constraints between successive signal values.)

If the clipped signal is subsequently differentiated [see Fig. 3(*c*)] the following expressions are obtained:

$$\mu_{xz} = 2^{\frac{3}{2}} A_x k^{\frac{1}{2}}/\sigma_x(1 + k) \quad . \quad . \quad . \quad . \quad (8)$$

and

$$\nu_{xz} = \frac{4}{\pi}\mu_{xz} \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

with k and A_x defined as before by eqns. (5) and (6). The numerical results are as follows:

$\mu_{xz} = 0 \cdot 635$, $\nu_{xz} = 0 \cdot 718$, for the Gaussian distribution,

$\mu_{xz} = 0 \cdot 500$, $\nu_{xz} = 0 \cdot 637$, for the exponential distribution.

Experimentally it is found[1–3] that subsequent differentiation has a negligible effect on intelligibility, which would seem to justify a preference for ν as a criterion of coherence in this case. However, it should be noted that, strictly speaking, eqns. (8) and (9) are valid only for totally random signals.

In addition to amplifying the above results, the paper is also concerned with the computation of the complete correlation and spectral-energy-density functions of both classes of signal. For example, it is shown that for a totally random signal subject to infinite clipping

$$\rho_{xx}(\tau) = \rho_{yy}(\tau) = \rho_{xy}(\tau)/\mu_{xy} \quad . \quad . \quad . \quad (10)$$

An attempt is made to ascertain under what conditions the same relationship would hold for partially constrained signals, but it would seem that no simple wholly exclusive criteria exist.[6] In this connection, it is interesting that preliminary experiments with a speech correlator[7] suggest a marked resemblance of $\rho_{xx}(\tau)$ and $\rho_{yy}(\tau)$ for a number of speech sounds.

In conclusion it must be emphasized that, although in the cases treated the coherence coefficients have numerical values approximating to measured articulation scores, no justification exists for assuming a formal relationship between coherence and intelligibility. However, it would seem likely that a first-order measure of similarity in spectral content must provide a rough indication of intelligibility, thus substantiating the high articulation scores observed experimentally.

References

(1) LICKLIDER, J. C. R., and POLLACK, I.: "Effect of Differentiation, Integration and Infinite Peak Clipping upon the Intelligibility of Speech," *Journal of the Acoustical Society of America*, 1948, **20**, p. 42.

(2) LICKLIDER, J. C. R.: "The Intelligibility of Amplitude-Dichotomised Time-Quantised Speech Waves," *ibid.*, 1950, **22**, p. 820.

(3) Unpublished work at the Imperial College of Science and Technology, London.

(4) JAMES, H. M., NICHOLS, N. B., and PHILLIPS, R. S.: "Theory of Servomechanisms," *M.I.T. Radiation Laboratory Series* No. 25 (McGraw-Hill, New York, 1947), Chapter 6.

(5) DAVENPORT, W. B.: "A Study of Speech Probability Distributions," *Journal of the Accoustical Society of America*, 1952, **24**, p. 390.

(6) LUCE, R. D.: "Amplitude Distorted Signals," *Quarterly Progress Report of the Research Laboratory of Electronics, M.I.T.*, April, 1953, p. 37.

(7) HOLMES, J. N., and DUKES, J. M. C.: "A Speech-Waveform Correlator with Magnetic-Tape Delay and Electronic Multiplication," *Proceedings I.E.E.*, Paper No. 1639 R, July, 1954 (**101**, Part III, p. 225).

THE GENERATION OF MILLIMETRE WAVES

621.372.2.029.65 Monograph No. 112 R

J. L. FARRANDS, Ph.D., B.Sc.

(DIGEST *of a paper published in December*, 1954, *as an* INSTITUTION MONOGRAPH *and republished in March*, 1955, *in Part C of the* PROCEEDINGS.)

That part of the electromagnetic spectrum lying between the highest radio frequencies and the far infra-red has provided a challenge to engineers and physicists for several decades. Experiments have been performed at various times[1,2,3] since 1923, and the sum of these may be said to have closed the gap with a limited degree of refinement. However, the subject is not systematized; the techniques used are not satisfactory for precise measurement or practical use.

It is natural that in problems of this kind a very wide variety of techniques may be attempted; for example, efforts have been made to use scaled models of conventional sources[4] and detectors, to extend the infra-red techniques,[3] to use Cerenkov radiation,[5] to induce Doppler radiation of high-speed electrons,[6] and to produce wide-band radiation from sparked resonators. It is only with the last of these that the paper is concerned in detail, although some reference is made to the first in view of its immediate practical importance. A systematic review of all the methods is beyond the compass of a single paper.

The sparked resonator is discussed in some detail in an attempt to discover what the power limitations are and whether any oscillation-maintaining effect is present in the spark. The special case of spherical resonators is treated because of the geometrical regularity and because appeal can be made to some of the most reliable experimental results.

It is concluded that the best explanation of these devices is that the inhomogeneous electrostatic charge induced by applied electric field redistributes itself over the surface in a periodic fashion following the collapse of the field on spark breakdown. From measurements of the spectral width it is shown that the damping of these oscillations is almost wholly attributable to radiation damping. This implies a high efficiency in the radiation process. There appears to be no evidence of an oscillation-maintaining effect in the spark.

Assumption of 100% radiation efficiency on these grounds permits calculation of available power in terms of the medium surrounding the spheres. It is concluded that spheres subjected to sparks in air are unlikely to produce useful amounts of power at short wavelengths, but that sparks in oil may be satisfactory.

An examination is made of some results reported using sparks in air. Evidence is deduced from these reports, including dielectric absorption measurements, which lead to some suspicions concerning their validity. It is also concluded that experiments with the mass radiator can be more conclusive if regular spherical particles are used.

Dr. Farrands is at the Defence Standards Laboratory, Australia, and was formerly at the Imperial College of Science and Technology.

REFERENCES

(1) McGUBBIN, JR., T. K., and SINTON, W. M.: "Far Infra-red Spectroscopy from 100 to 700 microns," *Journal of the Optical Society of America*, 1952, **42**, p. 113.
(2) NICHOLLS, E., and TEAR, J. D.: "Joining the Infra-red and Electric Wave Spectra," *Astrophysical Journal*, 1925, **61**, p. 17.
(3) GLAGOLEWA-ARKADIEWA, A. A.: "Short Electromagnetic Waves of Wavelength up to 82 μ," *Nature*, 1924, **113**, p. 610.
(4) LAFFERTY, J. M.: "A Millimetre Wave Reflex Oscillator," *Journal of Applied Physics*, 1946, **17**, p. 1061.
(5) GINSBURG, V. L.: "On the Utilisation of the Cherenkov Effect for the Radiation of Microwaves," *Comptes Rendues de l'U.R.S.S.*, 1947, **56**, p. 253.
(6) MOTZ, H.: "Applications of Radiation from Fast Electron Beams," *Journal of Applied Physics*, 1951, **22**, p. 527.

PROCEEDINGS OF THE INSTITUTION OF ELECTRICAL ENGINEERS

ISSUED IN THREE PARTS AS FOLLOWS:
Part A. POWER ENGINEERING (*February, April, etc.*)
Part B. RADIO AND ELECTRONIC ENGINEERING (*January, March, etc.*)
Part C. INSTITUTION MONOGRAPHS (*March and September only*)

Part B. RADIO AND ELECTRONIC ENGINEERING (INCLUDING COMMUNICATION ENGINEERING)

MARCH 1955

CONTENTS OF THIS ISSUE

Bibliographical References.—It is requested that bibliographical reference to an Institution paper should always include the serial number of the paper and the month and year of publication, which will be found at the top right-hand corner of the first page of the paper. This information should precede the reference to the Volume and Part.

Example.—SMITH, J.: "Reflections from the Ionosphere," *Proceedings I.E.E.*, Paper No. 3001 R, December, 1954 (102 B, p. 1234).

Published by The Institution, Savoy Place, London, W.C.2. *Telephone:* Temple Bar 7676. *Telegrams:* "Voltampere, Phone, London."
Printed by Unwin Brothers Limited, Woking and London.